Taking aim at cancer's master control genes p. 1174

Wind-borne microplastic pollution pp. 1184 & 1257

Actin actions in cell protrusion p. 1205

Science

$15
12 JUNE 2020

AAAS

CANARIES CHANGING COLORS

Mapping genes involved in sexual dimorphism pp. 1185 & 1270

READY FOR A TRULY LIFE-CHANGING MOMENT? APPLY NOW.

The *Science* & SciLifeLab Prize for Young Scientists is an annual prize awarded to early-career scientists. The prize is presented in four categories: Cell and Molecular Biology; Genomics, Proteomics, and Systems Biology Approaches; Ecology and Environment; and Molecular Medicine.

As a winner, you will have your essay published by *Science*, win up to USD 30,000 and be invited to Sweden where you will receive your award, present your research and meet with leading scientists in your field.

Get ready for a life-changing moment in your scientific career – apply now, deadline is July 15!

SCIENCEPRIZE.SCILIFELAB.SE

Knut och Alice Wallenbergs Stiftelse

Science AAAS

SciLifeLab

2019 winners (from left): Longzhi Tan, Zibo Chen, Barbara Klump and Humsa Venkatesh.

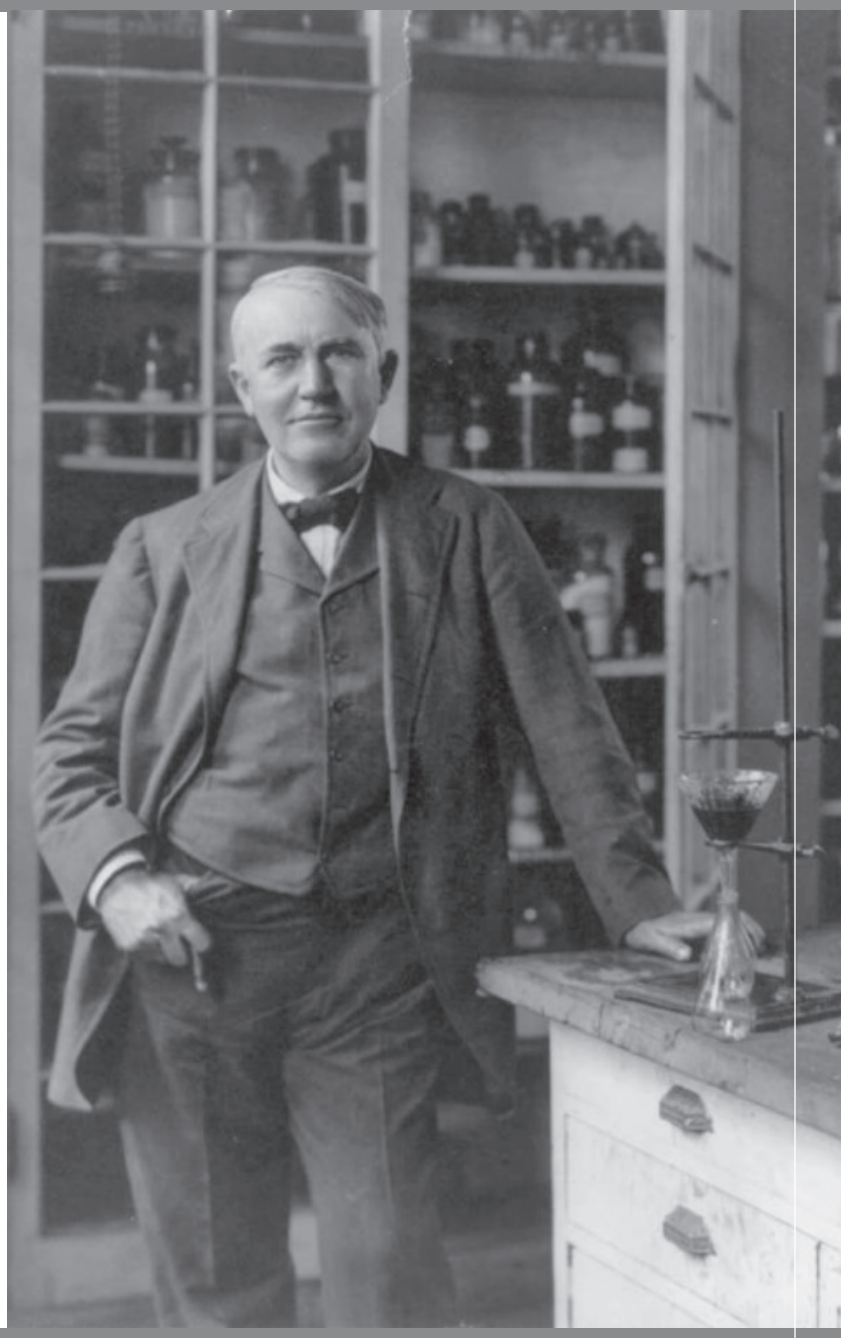
WHAT DO YOU AND THOMAS EDISON HAVE IN COMMON?

AAAS.

By investing in AAAS you join Thomas Edison and the many distinguished individuals whose vision led to the creation of AAAS and our world-renowned journal, *Science*, more than 150 years ago.

Like Edison, you can create a legacy that will last well into the future through planned giving to AAAS. By making AAAS a beneficiary of your will, trust, retirement plan, or life insurance policy, you make a strong investment in our ability to advance science in the service of society for years to come.

To discuss your legacy planning, contact Juli Staiano, Chief Philanthropy Officer, at (202) 326-6636, or jstaiano@aaas.org, or visit **aaas.org/1848society** for more information.



1848 SOCIETY AAAS

"I feel great knowing that I will leave behind a legacy that will be channeled through the AAAS. It also means a lot to me to be able to honor my late parents, too."

–PETER ECKEL
Member, 1848 Society and AAAS Member since 1988

AI—machines and software equipped with human-like intelligence—is now entering the drug discovery field.

Innovating AI today for better drug discovery tomorrow:

Machine Learning at Merck KGaA, Darmstadt, Germany, helps researchers find potential drugs faster and more easily

Artificial intelligence, or AI, has crept into our daily lives in many ways: via Siri who recognizes our voices, Alexa who tells jokes, and that helpful, unnamed woman who instructs us to turn left just as our cars approach an intersection. In the area of drug discovery, Merck KGaA, Darmstadt, Germany, is pushing AI to the next level to save lives, by using it to help find new and better treatments for disease—which it may be able to do more quickly and efficiently than is possible with traditional methods.

The type of AI we most frequently encounter—also known as Narrow AI—uses voice- or face-recognition programs (algorithms) that are trained using millions or billions of examples of voices or faces. In contrast, Merck KGaA, Darmstadt, Germany, is looking to use AI to solve problems in niche areas, such as drug discovery or materials science, where the data points for training the algorithms exist only in the tens of thousands.

"That makes it more challenging for us to build the AI models," says Helmut Linde, global head of Data Science & Analytics at Merck KGaA, Darmstadt, Germany. "While the most common machine-learning approaches rely on pattern detection in large data sets, we need to find new ways to combine this statistical approach with domain knowledge that we feed into the models explicitly."

In drug discovery, for example, Merck KGaA, Darmstadt, Germany, is already leading the way by using AI models and software to predict which chemical structures would make the best drug candidates to pursue, screen in the laboratory, and then test in preclinical and clinical trials. Only 10% of new drugs that enter into the first phase of clinical trials eventually get approved for use (*1*). AI experts and chemists at Merck KGaA, Darmstadt, Germany think that deploying machine-learning AI at the earliest stages of drug discovery—to identify the most promising molecular structures for drug candidates—will help improve that efficiency.

Machine learning uses neural networks modeled somewhat on the neural networks found in our brains. In machine learning, a computer model is first trained on a large set of data points or images that have first been sorted by humans. From this data set, the computer program "learns" all the characteristics that make up, for instance, an image of a panda. Once the training is over, the model should be able to then sift through new images on its own and correctly identify images of pandas, without any guidance from a human operator.

Machines learn to design drugs

Advanced machine learning, including deep learning, a special architecture of neural networks, can now be applied to drug discovery challenges. For example, the recent explosion in molecular knowledge about what causes disease has given us thousands of known drug targets—the molecules in the human body that are responsible for disease, which are mostly proteins. And we have billions of possible drug candidates—usually small, chemically synthesized molecules that could be discovered or designed and used to block, enhance, or modify these targets.

Identifying potential small-molecule drugs typically involves researchers either screening huge libraries of compounds to find those with a desired activity or designing novel compounds based on structures likely to produce that activity. Importantly, both approaches represent bottlenecks in drug discovery, because even with automated processes, it still takes 1–3 years and often the synthesis of a few thousand molecules to come up with a small molecule that modifies a target activity in the desired way and is selective for the desired target (to ensure there will be no side effects)—this represents a lot of money and person-hours.

AI systems offer very high data processing speed and efficiency that could accelerate the designing and screening process for drug candidates, resulting in the delivery of better treatments to patients more quickly. One study predicted that using AI in the drug discovery process could save USD 70 billion by 2028 (*2*).

"Predictive models are central to our work," says Friedrich Rippmann, director, Global Computational Chemistry & Biology at Merck KGaA, Darmstadt, Germany. "These are statistical models that predict whether a compound idea—a not-yet-synthesized molecule—will produce a desired activity."

These deep neural networks can identify novel compounds and the chemical groups responsible for wanted or unwanted activities. As Rippmann explains, the predictive models go further than simply informing scientists that "this compound will or will not work." The models can also tell chemists why something won't work and can

even highlight the areas of a molecule that are responsible for an activity. "This immediately indicates to chemists how to remove a certain unwanted activity from a candidate drug molecule," says Rippmann.

Rippman's team, working alongside company collaborators and partners, has developed about 300 new models to predict compound properties. The company's chemists use these models to decide which invented compounds have higher odds of success and should therefore be synthesized first for further testing.

SYNTHIA™ helps chemists walk backward

Organic chemists are the sophisticated experts who will synthesize the molecules suggested by Rippmann's group, based on the drug target protein as a starting point, and design them to inactivate it or modify it. In a practice called retrosynthesis, those chemists then work from the desired molecule backward to small starting materials to figure out the most feasible pathway to synthesize it.

MilliporeSigma, one of the three businesses of Merck KGaA, Darmstadt, Germany, is continuing to develop an AI-enabled software program called SYNTHIA™ to help chemists navigate retrosynthesis more efficiently. Coded by expert organic chemists and computer scientists over the course of more than 15 years, SYNTHIA™'s algorithms are driven by more than 100,000 expert-coded chemical reaction rules that describe known synthetic steps and allow the algorithms to stitch together several viable routes from commercial or known starting materials to the desired drug-like molecule.

At each retrosynthesis step, SYNTHIA™ considers local chemical interactions or potential conflicts between chemical groups within a molecule. SYNTHIA™ also explores both novel and known solutions at each step, eliminates unworkable options, and presents the synthesis pathways with the highest potential for success in the lab. To further refine the returned pathways, chemists can add filters—such as limiting the results to the fewest number of steps or the lowest cost for starting materials—making SYNTHIA™ a highly customized tool to fit the needs of the chemist.

However, there are limitations to current AI efforts. For computers, it is still a significant challenge to transfer knowledge from one domain to a new context, or even to make the knowledge it has "learned" understandable. An artificial neural network may be able to recognize cat breeds in images, but that does not mean that it can "explain" what a cat is or that it can use this knowledge (without human intervention) to get a head start in learning to identify dog breeds. These limitations have a particularly strong effect on the life sciences, where training data is often rare and expensive to obtain, while a huge corpus of domain knowledge is just waiting to be included in AI models.

Linde believes it will take more interdisciplinary work between neuroscientists, computer scientists, mathematicians, and theoretical physicists to find a new generation of machine-learning algorithms—algorithms that are inspired by a deeper understanding of the human brain. "The AI of the future must be much better at grasping the essence of things and deriving meaningful levels of abstraction on its own—rather than relying on a kind of interpolation between huge numbers of annotated data points," Linde says. If this hope is realized, researchers in the life sciences will find their digital helpers even more useful than today. And, who knows, maybe someday Alexa will even understand the jokes she tells us.

References

1. BIO, Biomedtracker, and Amplion, "Clinical Development Success Rates 2006–2015"; available at www.bio.org/sites/default/files/legacy/bioorg/docs/Clinical%20Development%20Success%20Rates%202006-2015%20-%20BIO,%20Biomedtracker,%20Amplion%202016.pdf.

2. K. Preuer, G. Klambauer, F. Rippmann, S. Hochreiter, T. Unterthiner, "Interpretable Deep Learning in Drug Discovery," in W. Samek, G. Montavon, A. Vedaldi, L. K. Hansen, K.-R. Müller, eds., *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, pp. 331–345, Lecture Notes in Computer Science, vol. 11700 (Springer, Cham, Switzerland, 2019).

For more info:

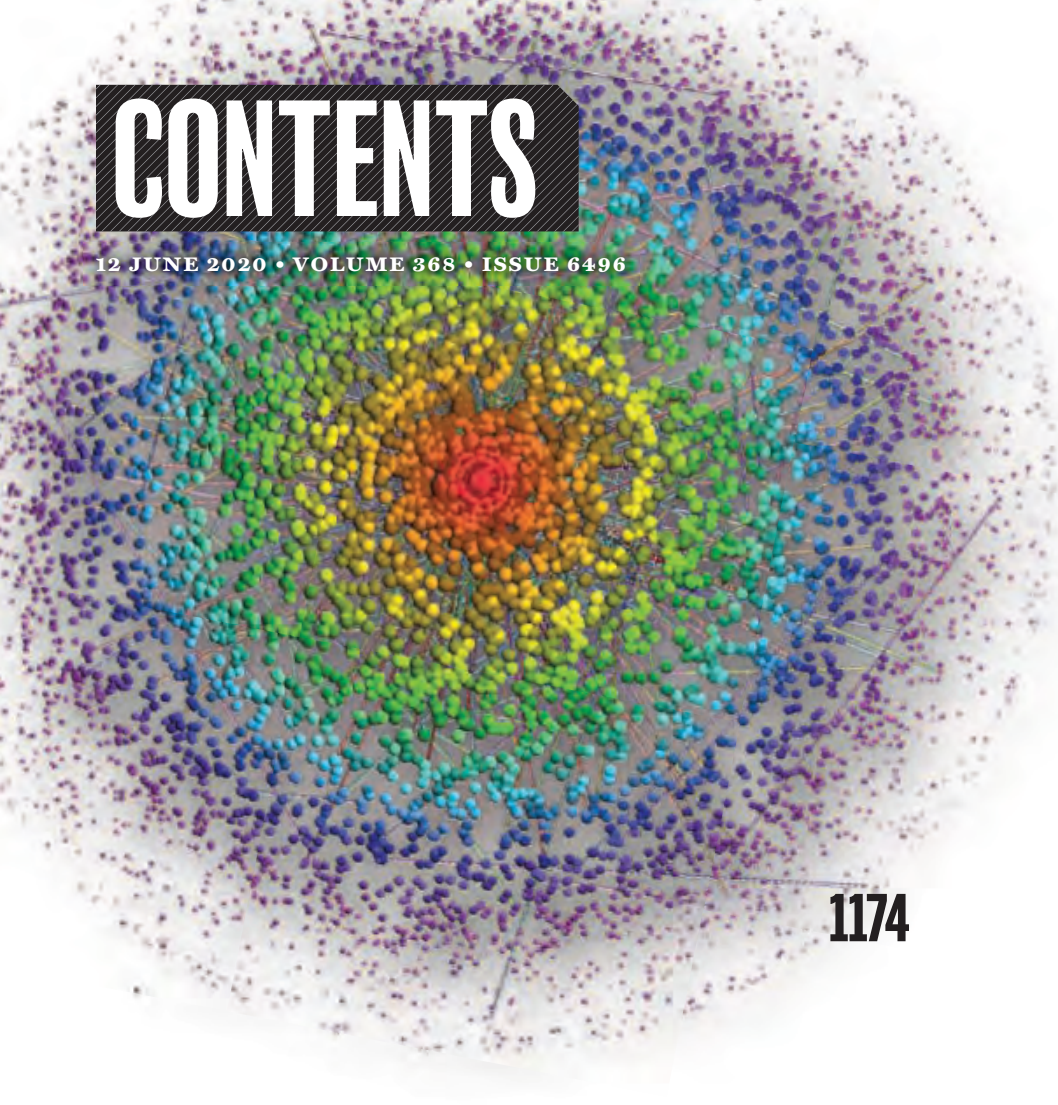
Worldwide (excl. USA and Canada):
www.merckgroup.com/en/research/science-space/envisioningtomorrow/smarter-connected-world/strongai.html

USA and Canada:
www.emdgroup.com/en/research/science-space/envisioning-tomorrow/smarter-connected-world/strongai.html

The company's AI-enabled software program, SYNTHIA™, now helps chemists more efficiently conduct retrosynthesis, the backwards creation of molecules to an already-known drug target protein.

CONTENTS

12 JUNE 2020 • VOLUME 368 • ISSUE 6496



1174

NEWS

IN BRIEF

1164 News at a glance

IN DEPTH

1166 Big studies dim hopes for hydroxychloroquine
Amid politicization and scandal, a disappointing scientific picture is emerging
By K. Kupferschmidt

1167 Authors, elite journals under fire after major retractions
Editors, co-authors missed warning signs, critics say *By C. Piller and J. Travis*

1169 Coronavirus rips through Dutch mink farms, triggering culls
Public concerns rise as two workers catch the disease *By M. Enserink*

1170 Vaccines that use human fetal cells draw fire
Abortion opponents urge United States and Canada to avoid "ethically-tainted" cell lines
By M. Wadman

1171 Galactic flash points to source of enigmatic fast radio bursts
CHIME telescope catches burst generated by a highly magnetized neutron star in the Milky Way *By D. Clery*

1172 Prominent Harvard anthropologist put on leave
Three former students accuse Inca scholar Gary Urton, alleging propositions or "traumatic" affairs *By A. Gibbons*

FEATURES

1174 Computing cancer's weak spots
An algorithm to unmask tumors' molecular linchpins is tested in patients *By R. Khamsi*

INSIGHTS

PERSPECTIVES

1178 Costs and benefits of living with predators
Integrating ecology and economics reveals a net monetary benefit from sea otters
By J. A. Estes and L. P. Carswell
REPORT p. 1243

1180 CRISPR at lightning speeds
Breaking the genome with the flick of a (light) switch *By D. Medhi and M. Jasin*
REPORT p. 1265

1181 Electrification of the chemical industry
Performing chemical synthesis with renewable electricity can reduce carbon emissions *By J. L. Barton*
REPORT p. 1228

1183 Mapping Earth's deepest secrets
Sequencing seismograms pinpoint new structures near Earth's core-mantle boundary *By M. S. Miller*
REPORT p. 1223

1184 The global odyssey of plastic pollution
Thinking big about small particles reveals new features of the microplastic cycle
By C. M. Rochman and T. Hoellein
REPORT p. 1257

1185 A gene for color differences between sexes
Sex differences in plumage color in hybrid canaries are controlled by a single enzyme *By N. Chen*
REPORT p. 1270

1187 Can existing live vaccines prevent COVID-19?
Live vaccines can prevent unrelated infections and may temporarily protect against COVID-19
By K. Chumakov et al.

1189 Lord Robert May (1936–2020)
Pioneering ecologist and a founder of chaos theory *By H. C. J. Godfray and A. R. McLean*

POLICY FORUM

1190 Moving academic research forward during COVID-19
A gradual, stepwise approach to reopening, informed by public health expertise, will be essential *By N. S. Wigginton* et al.

1193 A biodiversity target based on species extinctions
A single target comparable to the 2°C climate target may help galvanize biodiversity policy
By M. D. A. Rounsevell et al.

BOOKS *ET AL.*

1196 Discounting lives
Are COVID-19 deaths "sacrifices"? If so, to what? *By A. J. Wolfe*

1197 Labeling fish, labeling people
A science reporter grapples with an ichthyologist's all-consuming passion for categorization *By E. Landau*

IMAGE: SUMAZIN *ET AL.*, *CELL*, **147**, (2011)

LETTERS

1198 Saving the Sundarbans from development
By S. A. Mukul et al.

1198 Aggregating data from COVID-19 trials
By E. L. Ogburn et al.

1199 Brazil undermines parks by relocating staff
By P. R. Gonçalves et al.

RESEARCH

IN BRIEF

1201 From *Science* and other journals

REVIEW

1204 Neuroscience
Linking glutamate receptor movements and synapse function *L. Groc and D. Choquet*
REVIEW SUMMARY; FOR FULL TEXT: DX.DOI.ORG/10.1126/SCIENCE.AAY4631

RESEARCH ARTICLES

1205 Cell biology
Membrane-proximal F-actin restricts local membrane protrusions and directs cell migration *A. Bisaria* et al.

1211 Structural biology
Structures of cell wall arabinosyltransferases with the anti-tuberculosis drug ethambutol *L. Zhang* et al.

REPORTS

1219 Nanophotonics
Far-field excitation of single graphene plasmon cavities with ultracompressed mode volumes *I. Epstein* et al.

1223 Geophysics
Sequencing seismograms: A panoptic view of scattering in the core-mantle boundary region *D. Kim* et al.
PERSPECTIVE p. 1183

1228 Electrochemistry
Chloride-mediated selective electrosynthesis of ethylene and propylene oxides at high current density *W. R. Leow* et al.
PERSPECTIVE p. 1181

1234 Mesoscopic physics
h/e oscillations in interlayer transport of delafossites *C. Putzke* et al.

1239 Ice sheets
Pervasive ice sheet mass loss reflects competing ocean and atmosphere processes *B. Smith* et al.

1243 Ecological economics
Cascading social-ecological costs and benefits triggered by a recovering keystone predator *E. J. Gregr* et al.
PERSPECTIVE p. 1178; PODCAST

1247 Molecular machines
A precise polyrotaxane synthesizer *Y. Qiu* et al.

1253 Protein folding
Disordered proteins follow diverse transition paths as they fold and bind to a partner *J.-Y. Kim and H. S. Chung*

1257 Plastic pollution
Plastic rain in protected areas of the United States *J. Brahney* et al.
PERSPECTIVE p. 1184

1261 Cell cycle
Temporal integration of mitogen history in mother cells controls proliferation of daughter cells *M. Min* et al.

1265 Biotechnology
Very fast CRISPR on demand *Y. Liu* et al.
PERSPECTIVE p. 1180

1270 Sexual dimorphism
A genetic mechanism for sexual dichromatism in birds *M. A. Gazda* et al.
PERSPECTIVE p. 1185

1274 Coronavirus
A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2 *Y. Wu* et al.

DEPARTMENTS

1161 Editorial
Time to look in the mirror *By H. Holden Thorp*

1163 Editorial
COVID-19 and flu, a perfect storm
By Edward A. Belongia and Michael T. Osterholm

1282 Working Life
Helping others—and myself
By Moamen M. Elmassry

ON THE COVER

Male (left) and female (right) canaries (*Serinus canaria*) exhibiting sexual dichromatism—i.e., color difference between sexes. These *mosaic* canaries were created by crossing common canaries with red siskins and selecting the progeny for sexual dichromatism. The pronounced differences in plumage coloration are driven by a single gene that mediates sexual dichromatism in birds. See pages 1185 and 1270. *Photo: Ricardo Jorge Lopes*

Science Staff 1160
Science Careers 1279

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices.

Copyright © 2020 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS.

Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2148; Foreign postage extra: Air assist delivery: $98. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling available from backissues.sciencemag.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

PHOTO: JAMES THOMPSON PHOTOGRAPHY

Science AAAS

Editor-in-Chief Holden Thorp, hthorp@aaas.org

Executive Editor Monica M. Bradford

Editors, Research Valda Vinson, Jake S. Yeston **Editor, Insights** Lisa D. Chong

DEPUTY EDITORS Julia Fahrenkamp-Uppenbrink (UK), Stella M. Hurtley (UK), Phillip D. Szuromi, Sacha Vignieri **SR. EDITORIAL FELLOW** Andrew M. Sugden (UK) **SR. EDITORS** Gemma Alderton (UK), Caroline Ash (UK), Brent Grocholski, Pamela J. Hines, Marc S. Lavine (Canada), Ian S. Osborne (UK), Beverly A. Purnell, L. Bryan Ray, H. Jesse Smith, Keith T. Smith (UK), Jelena Stajic, Peter Stern (UK), Valerie B. Thompson, Brad Wible, Laura M. Zahn **ASSOCIATE EDITORS** Michael A. Funk, Priscilla N. Kelly, Tage S. Rai, Seth Thomas Scanlon (UK), Yury V. Suleymanov **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Harry Jach, Lauren Kmec **CONTENT PRODUCTION EDITORS** Amelia Beyna, Jeffrey E. Cook, Chris Filiatreau, Julia Katris, Nida Masiulis, Suzanne M. White **SR. EDITORIAL COORDINATORS** Carolyn Kyle, Beverly Shields **EDITORIAL COORDINATORS** Aneera Dobbins, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Maryrose Madrid, Ope Martins, Shannon McMahon, Jerry Richardson, Hilary Stewart (UK), Alana Warnke, Alice Whaley (UK), Anita Wynn **PUBLICATIONS ASSISTANTS** Jeremy Dow, Alexander Kief, Ronmel Navas, Brian White **EXECUTIVE ASSISTANT** Jessica Slater **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI SR. OFFICE ADMINISTRATOR** Jessica Waldock (UK)

News Editor Tim Appenzeller

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** Martin Enserink **DEPUTY NEWS EDITORS** Elizabeth Culotta, Lila Guterman, David Grimm, Eric Hand (Europe), David Malakoff **SR. CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brainard, Catherine Matacic **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Kelly Servick, Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **INTERNS** Eva Frederick, Rodrigo Perez Ortega **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Ann Gibbons, Mara Hvistendahl, Sam Kean, Eli Kintisch, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Eliot Marshall, Virginia Morell, Dennis Normile (Shanghai), Elisabeth Pain (Careers), Charles Piller, Michael Price, Tania Rabesandratana (Barcelona), Emily Underwood, Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Donisha Adams, Rachel Bernstein (Editor), Katie Langin (Acting Editor) **COPY EDITORS** Julia Cole (Senior Copy Editor), Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

Creative Director Beth Rakouskas

DESIGN MANAGING EDITOR Marcy Atarod **GRAPHICS MANAGING EDITOR** Alberto Cuadra **PHOTOGRAPHY MANAGING EDITOR** William Douthitt **WEB CONTENT STRATEGY MANAGER** Kara Estelle-Powers **DESIGN EDITOR** Chrystal Smith **DESIGNER** Christina Aycock **GRAPHICS EDITOR** Nirja Desai **INTERACTIVE GRAPHICS EDITOR** Xing Liu **SENIOR SCIENTIFIC ILLUSTRATORS** Valerie Altounian, Chris Bickel **SCIENTIFIC ILLUSTRATOR** Alice Kitterman **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **SENIOR PHOTO EDITOR** Emily Petersen **PHOTO EDITOR** Kaitlyn Dolan

Chief Executive Officer and Executive Publisher Sudip Parikh

Publisher, Science Family of Journals Bill Moran

DIRECTOR, BUSINESS SYSTEMS AND FINANCIAL ANALYSIS Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYSIS** Eric Knott **DIRECTOR OF ANALYTICS** Enrique Gonzales **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **SENIOR BUSINESS ANALYST** Cory Lipman, Meron Kebede **FINANCIAL ANALYST** Alexander Lee **ADVERTISING SYSTEM ADMINISTRATOR** Tina Burks **SENIOR SALES COORDINATOR** Shirley Young **DIGITAL/PRINT STRATEGY MANAGER** Jason Hillman **QUALITY TECHNICAL MANAGER** Marcus Spiegler **ASSISTANT MANAGER DIGITAL/PRINT** Rebecca Doshi **SENIOR CONTENT SPECIALISTS** Steve Forrester, Jacob Hedrick, Antoinette Hodal, Lori Murphy **DIGITAL PRODUCTION MANAGER** Lisa Stanford **CONTENT SPECIALIST** Kimberley Oster **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **DESIGNER, CUSTOM PUBLISHING** Jeremy Huntsinger **SR. TRAFFIC ASSOCIATE** Christine Hall **SPECIAL PROJECTS ASSOCIATE** Sarah Dhere

ASSOCIATE DIRECTOR, BUSINESS DEVELOPMENT Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **DIGITAL MARKETING MANAGER** Aimee Aponte **JOURNALS MARKETING MANAGER** Shawana Arnold **MARKETING ASSOCIATES** Tori Velasquez, Mike Romano, Ashley Hylton **DIGITAL MARKETING SPECIALIST** Asleigh Rojanavongse **SENIOR DESIGNER** Kim Huynh

DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING Sean Sanders **ASSISTANT EDITOR, CUSTOM PUBLISHING** Jackie Oberst

DIRECTOR, PRODUCT & PUBLISHING DEVELOPMENT Chris Reid **DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT** Sarah Whalen **ASSOCIATE DIRECTOR, PRODUCT MANAGMENT** Kris Bishop **SR. PRODUCT ASSOCIATE** Robert Koepke **DIGITAL PRODUCT STRATEGIST** Michael Hardesty **SPJ ASSOCIATE** Samantha Bruno Fuller

DIRECTOR, INSTITUTIONAL LICENSING Iquo Edim **ASSOCIATE DIRECTOR, RESEARCH & DEVELOPMENT** Elisabeth Leonard **MARKETING MANAGER** Kess Knight **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Marco Castellan **MANAGER, AGENT RELATIONS & CUSTOMER SUCCESS** Judy Lillibridge **SENIOR OPERATIONS ANALYST** Lana Guz **FULFILLMENT COORDINATOR** Melody Stringer **SALES COORDINATOR** Josh Haverlock

DIRECTOR, GLOBAL SALES Tracy Holmes **US EAST COAST AND MID WEST SALES** Stephanie O'Connor **US WEST COAST SALES** Lynne Stickrod **US SALES MANAGER, SCIENCE CAREERS** Claudia Paulsen-Young **US SALES REP, SCIENCE CAREERS** Tracy Anderson **ASSOCIATE DIRECTOR, ROW** Roger Goncalves **SALES REP, ROW** Sarah Lelarge **SALES ADMIN ASSISTANT, ROW** Bryony Cousins **DIRECTOR OF GLOBAL COLLABORATION AND ACADEMIC PUBLISHING RELATIONS, ASIA** Xiaoying Chu **ASSOCIATE DIRECTOR, INTERNATIONAL COLLABORATION** Grace Yao **SALES MANAGER** Danny Zhao **MARKETING MANAGER** Kilo Lan **ASCA CORPORATION, JAPAN** Kaoru Sasaki (Tokyo), Miyuki Tani (Osaka) **COLLABORATION/CUSTOM PUBLICATIONS/JAPAN** Adarsh Sandhu

DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS Emilie David **RIGHTS AND LICENSING COORDINATOR** Jessica Adams **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **CONTRACTS AND LICENSING ASSOCIATE** Lili Catlett

---

MAIN HEADQUARTERS
Science/AAAS
1200 New York Ave. NW
Washington, DC 20005

SCIENCE INTERNATIONAL
Clarendon House
Clarendon Road
Cambridge, CB2 8FH, UK

SCIENCE CHINA
Room 1004, Culture Square
No. 59 Zhongguancun St.
Haidian District, Beijing, 100872

SCIENCE JAPAN
ASCA Corporation
Sibaura TY Bldg. 4F, 1-14-5
Shibaura Minato-ku
Tokyo, 108-0073 Japan

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

INFORMATION FOR AUTHORS
sciencemag.org/authors/science-information-authors

REPRINTS AND PERMISSIONS
sciencemag.org/help/reprints-and-permissions

MEDIA CONTACTS
scipak@aaas.org

MULTIMEDIA CONTACTS
SciencePodcast@aaas.org
ScienceVideo@aaas.org

INSTITUTIONAL SALES AND SITE LICENSES
sciencemag.org/librarian

PRODUCT ADVERTISING & CUSTOM PUBLISHING
advertising.sciencemag.org/products-services
science_advertising@aaas.org

CLASSIFIED ADVERTISING
advertising.sciencemag.org/science-careers
advertise@sciencecareers.org

JOB POSTING CUSTOMER SERVICE
employers.sciencecareers.org
support@sciencecareers.org

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
sciencemag.org/subscriptions

MEMBER BENEFITS
aaas.org/membercentral

AAAS BOARD OF DIRECTORS
CHAIR Steven Chu
PRESIDENT Claire M. Fraser
PRESIDENT-ELECT Susan G. Amara
TREASURER Carolyn N. Ainslie
CHIEF EXECUTIVE OFFICER
Sudip Parikh
BOARD Cynthia M. Beall
Rosina M. Bierbaum
Ann Bostrom
Stephen P.A. Fodor
S. James Gates, Jr.
Laura H. Greene
Kaye Husbands Fealing
Maria M. Klawe
Robert B. Millard
Alondra Nelson
William D. Provine

BOARD OF REVIEWING EDITORS (Statistics board members indicated with **(S)**)

Adriano Aguzzi, *U. Hospital Zürich*
Takuzo Aida, *U. of Tokyo*
Leslie Aiello, *Wenner-Gren Foundation*
Judith Allen, *U. of Manchester*
Sebastian Amigorena, *Institut Curie*
James Analytis, *U. of California, Berkeley*
Paola Arlotta, *Harvard U.*
Johan Auwerx, *EPFL*
David Awschalom, *U. of Chicago*
Clare Baker, *U. of Cambridge*
Nenad Ban, *ETH Zürich*
Franz Bauer, *Pontificia Universidad Católica de Chile*
Ray H. Baughman, *U. of Texas at Dallas*
Peter Bearman, *Columbia U.*
Carlo Beenakker, *Leiden U.*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke U.*
Gabriele Bergers, *VIB*
Bradley Bernstein, *Mass. General Hospital*
Alessandra Biffi, *Harvard Med. School*
Peer Bork, *EMBL*
Chris Bowler, *École Normale Supérieure*
Ian Boyd, *U. of St. Andrews*
Emily Brodsky, *U. of California, Santa Cruz*
Ron Brookmeyer, *U. of California, Los Angeles* **(S)**
Christian Büchel, *UKE Hamburg*
Dennis Burton, *Scripps Research*
Carter Tribley Butts, *U. of California, Irvine*
György Buzsáki, *New York U. School of Med.*
Blanche Capel, *Duke U.*
Annmarie Carlton, *U. of California, Irvine*
Nick Chater, *U. of Warwick*
Zhijian Chen, *UT Southwestern Med. Ctr.*
Ib Chorkendorff, *Denmark TU*
James J. Collins, *MIT*
Robert Cook-Deegan, *Arizona State U.*
Alan Cowman, *Walter & Eliza Hall Inst.*
Carolyn Coyne, *U. of Pittsburgh*
Roberta Croce, *VU Amsterdam*
Jeff L. Dangl, *U. of North Carolina*
Tom Daniel, *U. of Washington*
Chiara Daraio, *Caltech*
Nicolas Dauphas, *U. of Chicago*
Frans de Waal, *Emory U.*
Claude Desplan, *New York U.*
Sandra Díaz, *Universidad Nacional de Córdoba*
Ulrike Diebold, *TU Wien*
Hong Ding, *Inst. of Physics, CAS*
Jennifer Dionne, *Stanford U.*
Dennis Discher, *U. of Penn.*
Gerald Dorn, *Washington U. in St. Louis*
Jennifer A. Doudna, *U. of California, Berkeley*
Bruce Dunn, *U. of California, Los Angeles*
William Dunphy, *Caltech*
Christopher Dye, *U. of Oxford*
Todd Ehlers, *U. of Tübingen*
Jennifer Elisseeff, *Johns Hopkins U.*
Tim Elston, *U. of North Carolina*
Andrea Encalada, *U. San Francisco de Quito*
Nader Engheta, *U. of Penn.*
Karen Ersche, *U. of Cambridge*
Barry Everitt, *U. of Cambridge*
Vanessa Ezenwa, *U. of Georgia*
Michael Feuer, *The George Washington U.*
Toren Finkel, *U. of Pittsburgh Med. Ctr.*
Gwenn Flowers, *Simon Fraser U.*
Peter Fratzl, *Max Planck Inst. Potsdam*
Elaine Fuchs, *Rockefeller U.*
Eileen Furlong, *EMBL*
Jay Gallagher, *U. of Wisconsin*
Daniel Geschwind, *U. of California, Los Angeles*
Karl-Heinz Glassmeier, *TU Braunschweig*
Ramon Gonzalez, *U. of South Florida*
Elizabeth Grove, *U. of Chicago*
Nicolas Gruber, *ETH Zürich*
Hua Guo, *U. of New Mexico*
Kip Guy, *U. of Kentucky College of Pharmacy*
Taekjip Ha, *Johns Hopkins U.*
Christian Haass, *Ludwig Maximilians U.*
Sharon Hammes-Schiffer, *Yale U.*
Wolf-Dietrich Hardt, *ETH Zürich*
Louise Harra, *U. College London*
Jian He, *Clemson U.*
Carl-Philipp Heisenberg, *IST Austria*
Ykä Helariutta, *U. of Cambridge*
Janet G. Hering, *Eawag*
Hans Hilgenkamp, *U. of Twente*
Kai-Uwe Hinrichs, *U. of Bremen*
Lora Hooper, *UT Southwestern Med. Ctr.*
Fred Hughson, *Princeton U.*
Randall Hulet, *Rice U.*
Auke Ijspeert, *EPFL*
Akiko Iwasaki, *Yale U.*
Stephen Jackson, *USGS and U. of Arizona*
Kai Johnsson, *EPFL*
Peter Jonas, *IST Austria*
Matt Kaeberlein, *U. of Washington*
William Kaelin Jr., *Dana-Farber Cancer Inst.*
Daniel Kammen, *U. of California, Berkeley*
V. Narry Kim, *Seoul Nat. U.*
Robert Kingston, *Harvard Med. School*
Nancy Knowlton, *Smithsonian Institution*
Etienne Koechlin, *École Normale Supérieure*
Alex L. Kolodkin, *Johns Hopkins U.*
Julija Krupic, *U. of Cambridge*
Thomas Langer, *Max Planck Inst. Cologne*
Mitchell A. Lazar, *U. of Penn.*
Ottoline Leyser, *U. of Cambridge*
Wendell Lim, *U. of California, San Francisco*
Jianguo Liu, *Michigan State U.*
Luis Liz-Marzán, *CIC biomaGUNE*
Jonathan Losos, *Washington U. in St. Louis*
Ke Lu, *Chinese Acad. of Sciences*
Christian Lüscher, *U. of Geneva*
Jean Lynch-Stieglitz, *Georgia Inst. of Tech.*
Fabienne Mackay, *QIMR Berghofer*
Anne Magurran, *U. of St. Andrews*
Oscar Marín, *King's College London*
Charles Marshall, *U. of California, Berkeley*
Christopher Marx, *U. of Idaho*
Geraldine Masson, *CNRS*
C. Robertson McClung, *Dartmouth College*
Rodrigo Medellín, *U. Nacional Autónoma de México*
Graham Medley, *London School of Hygiene & Tropical Med.*
Jane Memmott, *U. of Bristol*
Baoxia Mi, *U. of California, Berkeley*
Edward Miguel, *U. of California, Berkeley*
Tom Misteli, *NCI, NIH*
Yasushi Miyashita, *U. of Tokyo*
Alison Motsinger-Reif, *NIEHS, NIH* **(S)**
Daniel Nettle, *Newcastle U.*
Daniel Neumark, *U. of California, Berkeley*
Beatriz Noheda, *U. of Groningen*
Helga Nowotny, *Vienna Science, Research & Tech. Fund*
Rachel O'Reilly, *U. of Birmingham*
Harry Orr, *U. of Minnesota*
Pilar Ossorio, *U. of Wisconsin*
Andrew Oswald, *U. of Warwick*
Isabella Pagano, *Istituto Nazionale di Astrofisica*
Margaret Palmer, *U. of Maryland*
Elizabeth Levy Paluck, *Princeton U.*
Jane Parker, *Max Planck Inst. Cologne*
Giovanni Parmigiani, *Dana-Farber Cancer Inst.* **(S)**
Samuel Pfaff, *Salk Inst. for Biological Studies*
Julie Pfeiffer, *UT Southwestern Med. Ctr.*
Matthieu Piel, *Institut Curie*
Kathrin Plath, *U. of California, Los Angeles*
Martin Plenio, *Ulm U.*
Katherine Pollard, *U. of California, San Francisco*
Elvira Poloczanska, *Alfred-Wegener-Inst.*
Julia Pongratz, *Ludwig Maximilians U.*
Philippe Poulin, *CNRS*
Jonathan Pritchard, *Stanford U.*
Félix A. Rey, *Institut Pasteur*
Trevor Robbins, *U. of Cambridge*
Joeri Rogelj, *Imperial College London*
Amy Rosenzweig, *Northwestern U.*
Mike Ryan, *U. of Texas at Austin*
Mitinori Saitou, *Kyoto U.*
Shimon Sakaguchi, *Osaka U.*
Miquel Salmeron, *Lawrence Berkeley Nat. Lab*
Nitin Samarth, *Penn. State U.*
Jürgen Sandkühler, *Med. U. of Vienna*
Alexander Schier, *Harvard U.*
Wolfram Schlenker, *Columbia U.*
Susannah Scott, *U. of California, Santa Barbara*
Rebecca Sear, *London School of Hygiene & Tropical Med.*
Vladimir Shalaev, *Purdue U.*
Jie Shan, *Cornell U.*
Beth Shapiro, *U. of California, Santa Cruz*
Jay Shendure, *U. of Washington*
Steve Sherwood, *U. of New South Wales*
Brian Shoichet, *U. of California, San Francisco*
Robert Siliciano, *Johns Hopkins U. School of Med.*
Lucia Sivilotti, *U. College London*
Alison Smith, *John Innes Centre*
Richard Smith, *U. of North Carolina* **(S)**
Mark Smyth, *QIMR Berghofer*
Pam Soltis, *U. of Florida*
John Speakman, *U. of Aberdeen*
Tara Spires-Jones, *U. of Edinburgh*
Allan C. Spradling, *Carnegie Institution for Science*
V. S. Subrahmanian, *Dartmouth College*
Ira Tabas, *Columbia U.*
Sarah Teichmann, *Wellcome Sanger Inst.*
Rocio Titiunik, *Princeton U.*
Shubha Tole, *Tata Inst. of Fundamental Research*
Wim van der Putten, *Netherlands Inst. of Ecology*
Reinhilde Veugelers, *KU Leuven*
Bert Vogelstein, *Johns Hopkins U.*
Kathleen Vohs, *U. of Minnesota*
David Wallach, *Weizmann Inst. of Science*
Jane-Ling Wang, *U. of California, Davis* **(S)**
David Waxman, *Fudan U.*
Jonathan Weissman, *U. of California, San Francisco*
Chris Wikle, *U. of Missouri* **(S)**
Terrie Williams, *U. of California, Santa Cruz*
Ian A. Wilson, *Scripps Research* **(S)**
Yu Xie, *Princeton U.*
Jan Zaanen, *Leiden U.*
Kenneth Zaret, *U. of Penn. School of Med.*
Jonathan Zehr, *U. of California, Santa Cruz*
Xiaowei Zhuang, *Harvard U.*
Maria Zuber, *MIT*

Science serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority or conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

EDITORIAL

Time to look in the mirror

H. Holden Thorp
Editor-in-Chief, *Science* journals.
hthorp@aaas.org;
@hholdenthorp

This is a grave time in American history. Both the public health and economic problems of coronavirus disease 2019 (COVID-19) were foreseeable. But even more predictable is the racial tension gripping the United States in the wake of the brutal killing of George Floyd. It is easy to think that the problem is isolated to individual racists in the community and in the government, and that the scientific enterprise is immune to racism. Scientific inquiry produces knowledge, and that ultimately leads to justice, right?

Not so fast. The U.S. scientific enterprise is predominantly white, as are the U.S. institutions that *Science*'s authors are affiliated with. The evidence of systemic racism in science permeates this nation. Why are so few *Science* authors from historically black colleges and universities? Why are the scientific areas studied more frequently by people of color continuously underfunded by the government? Why do students who are people of color have to remind society that they are almost never taught by someone who looks like them? Why has the United States failed to update its ways of teaching science when data show that people of color learn better with more inclusive methods? If there had been more diversity in science, would we have the painful legacy of the Tuskegee syphilis study and the shameful nonrecognition of Henrietta Lacks's contribution to science?

"...scientists finally need to listen to, and make space for, people of color to lead..."

Dr. Lisa White, a professor at the University of California, Berkeley, and chair of the American Geophysical Union's Diversity and Inclusion Advisory Committee, pointed out recently that environmental racism wouldn't be such a problem if there were a more diverse science professoriate. For example, only 4% of tenured and tenure track faculty in the top 100 geoscience departments in the United States are people of color.

Not surprisingly, Dr. Martin Luther King Jr. described this problem in 1963 in his "Letter from a Birmingham Jail":

"First, I must confess that over the past few years I have been gravely disappointed with the white moderate. I have almost reached the regrettable conclusion that the Negro's great stumbling block in his stride toward freedom is not the White Citizen's Counciler or the Ku Klux Klanner, but the white moderate, who is more devoted to 'order' than to justice; who prefers a negative peace which is the absence of tension to a positive peace which is the presence of justice; who constantly says: 'I agree with you in the goal you seek, but I cannot agree with your methods of direct action'; who paternalistically believes he can set the timetable for another man's freedom; who lives by a mythical concept of time and who constantly advises the Negro to wait for a 'more convenient season.'"

The reckoning Dr. King calls for has not happened in the intervening 57 years. The failure of the white moderates to heed the call of the Birmingham Jail is just as integral to today's systemic racism as the racist actions of some law enforcement. It's not just abusive police that need to be reminded that Black Lives Matter.

It is time for the scientific establishment to confront this reality and to admit its role in perpetuating it. The first step is for science and scientists to say out loud that they have benefited from, and failed to acknowledge, white supremacy. And then science and scientists finally need to listen to, and make space for, people of color to lead laboratories that publish great science and produce influential scientists, run institutions and their scientific units, and propel *Science* and other journals to promote structurally underfunded scientists and areas of science.

Someone I turn to for wisdom and leadership on this issue is Dr. Valerie Sheares Ashby, the Dean of the Trinity College of Arts and Sciences at Duke University. As an African-American woman, she claims that she is leading today only because a few people decided to let her into this exclusive club—something she says hardly ever happens. "How much creativity are we leaving on the table," she asks, "because science repeatedly fails to come to terms with our narrowly defined processes and our limited ways of determining success?"

As in the past, the scientific community is expressing anguish, outrage, and renewed commitment to promote equity and inclusion. But when the protests wind down and disappear from the headlines, science will be at a familiar fork in the road. Let's have the courage to take the right path this time.

–H. Holden Thorp

PHOTO: CAMERON DAVIDSON

Published online 8 June 2020; 10.1126/science.abd1896

THE BRAIN PRIZE

NOMINATION OF THE WORLD'S BEST BRAIN RESEARCHER 2021

Nominations must be received by 1 September 2020

The DKK 10 million is awarded to one or more scientists from around the world in recognition of their outstanding contribution to neuroscience – from basic to clinical.
Open for nominations now. Diversity among the nominees is encouraged.
For more details please visit **lundbeckfoundation.com/thebrainprize**

THE BRAIN PRIZE SELECTION COMMITTEE

Richard Morris, United Kingdom – Chairman
Story Landis, USA – Vice-Chairwoman
Catherine Dulac, USA
Geoffrey Donnan, Australia
Irene Tracey, United Kingdom
Joseph Coyle, USA
Mu-ming Poo, China
Ole Petter Ottersen, Sweden
Philip Scheltens, The Netherlands

LUNDBECK FOUNDATION

EDITORIAL

COVID-19 and flu, a perfect storm

Edward A. Belongia is the director of the Center for Clinical Epidemiology and Population Health, Marshfield Clinic Research Institute, Marshfield, WI, USA. belongia.edward@marshfieldclinic.org

Michael T. Osterholm is the director of the Center for Infectious Disease Research and Policy, University of Minnesota, Minneapolis, MN, USA. mto@umn.edu

The world is in uncharted waters for the 2020 respiratory virus season. For the first time in modern history, the Northern Hemisphere faces the prospect of the coronavirus disease 2019 (COVID-19) pandemic and a simultaneous epidemic of seasonal influenza. Each causes life-threatening illness and death, especially in older adults, people with chronic diseases, and other vulnerable populations. How can we prepare for this convergence?

The timing and severity of a COVID-19 wave in the fall and winter are uncertain, but past experiences with the 1918 and 1957 influenza pandemics point to the possibility of a resurgence. Almost nothing is known about the interaction of influenza virus and severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2, the cause of COVID-19) within individuals. Does coinfection increase the risk of severe illness or amplify virus shedding? Few coinfections have been reported from China during the early phase of the pandemic. The Southern Hemisphere influenza season is just beginning, and it may provide some clues as to what can be expected in the Northern Hemisphere later this year.

> **"How can we prepare for this convergence?"**

Much of the population remains susceptible to SARS-CoV-2, and the stress on hospitals will be greatest if the COVID-19 and influenza epidemics overlap and peak around the same time. It is possible that the number of individuals infected with each virus will peak at different times, reducing the peak demand for hospital beds. If a surge in COVID-19 cases occurs this fall, tightening mitigation strategies will be necessary. Social distancing and stay-at-home orders are socially and economically disruptive, but can reduce demand on hospitals and protect vulnerable populations. They will also reduce transmission of other respiratory viruses, including influenza and respiratory syncytial virus. Supplies of personal protective equipment must sufficiently meet the projected demand of a severe influenza season along with COVID-19.

There are important differences in the epidemiology of COVID-19 and seasonal influenza, but symptoms overlap. Molecular diagnostic testing for SARS-CoV-2 is critical for all patients with acute respiratory illness, especially during periods of co-circulation. Rapid-turnaround testing is necessary to distinguish between influenza and COVID-19, guide patient care, and support a comprehensive COVID-19 control program (including isolation of cases and rapid identification and quarantine of contacts). Every effort should be made to ensure that resources will be available for combined testing for COVID-19 and influenza. These tests should be without charge to patients because they serve a dual purpose for public health and patient care.

We do not yet have a COVID-19 vaccine, but safe and moderately effective influenza vaccines are available. Their widespread use is more important now than ever, and we encourage health care providers, employers, and community leaders to promote vaccination. Vaccine effectiveness varies by season and subtype, but vaccination offers similar protection against laboratory-confirmed influenza hospitalization and outpatient illness. Widespread misinformation on social media includes the false claim that influenza vaccination increases the risk of SARS-CoV-2 infection. Scientists, health care providers, and public health leaders must counter these claims with clear, evidence-based information on the importance of influenza vaccination during the COVID-19 pandemic.

The prospect of a second COVID-19 wave requires planning to ensure optimal delivery of influenza vaccines starting in the early fall. Community vaccination sites are often set up for maximum volume and efficiency, and alternative approaches will be needed to maintain physical distancing and minimize the risk of SARS-CoV-2 transmission, particularly because many influenza vaccine recipients are at high risk for both influenza and COVID-19 complications. Randomized trials have shown that some enhanced influenza vaccines (e.g., high-dose inactivated vaccine and recombinant vaccines) generate greater protection in older adults compared to standard-dose inactivated vaccine. The optimal timing of influenza vaccination in patients with confirmed COVID-19 is uncertain. There are no clinical studies on the effects of influenza vaccination in patients with COVID-19, but it may be prudent to delay vaccine administration until after the acute illness has resolved.

Over 400,000 COVID-19 deaths were reported worldwide by 6 June, including over 109,000 in the United States. The actual death count is almost certainly higher, and we are still in the early phase of the pandemic. The U.S. Centers for Disease Control and Prevention estimates that influenza has killed 12,000 to 61,000 each season over the past decade. Will there be a perfect storm of COVID-19 and influenza during the 2020–2021 season? We do not yet know, but we must start preparing in the coming months.

–Edward A. Belongia and Michael T. Osterholm

10.1126/science.abd2220

NEWS

Perhaps partly because mosques reopened in May, Iran has seen its COVID-19 case numbers rise rapidly this month.

IN BRIEF

Edited by **Lila Guterman**

DISPATCHES FROM THE PANDEMIC

Iran's infection numbers climb

PUBLIC HEALTH | Iran may be the first nation to experience a second wave of COVID-19. After a monthlong lull in cases, its health ministry reported a record 3574 new infections on 4 June. No country in the Middle East has been hit harder than Iran, which, by 9 June, had tallied 175,927 lab-confirmed cases and 8425 deaths. Its first wave peaked around 1 April, hammering the country's north, especially Tehran and the holy city of Qom. COVID-19 hot spots are now flaring in the south, with about one-quarter of recent cases reported in oil-rich Khuzestan province. Following a lockdown in April, mosques and schools reopened in mid-May, and domestic travel increased. Iran is also looking harder for COVID-19 with expanded testing, says Mohammad Saeid Rezaee-Zavareh, a physician at Sarallah Hospital in Karaj. "Only time can determine," he says, "whether it is a second wave or not."

Wuhan declared free of COVID-19

TESTING | After a marathon testing effort, Wuhan, China, the city where the coronavirus pandemic seems to have originated, pronounced itself free of the disease last week—at least under the definition used in China. Organized by the local and provincial governments, teams tested 9.9 million people between 14 May and 1 June. For efficiency they pooled swabs from five people at a time to test together. Officials found and isolated 300 people who carried the virus but had no symptoms; China does not count asymptomatic people as cases. No positive tests were found among more than 1100 contacts of the asymptomatic people. Teams also found no signs of the virus on the asymptomatic people's toothbrushes, mobile phones, and door handles.

Cow antibodies versus coronavirus

EXPERIMENTAL TREATMENTS | A biotech company has coaxed genetically modified cows to pump out human antibodies that subdue SARS-CoV-2, and it plans to start clinical trials this summer. SAb Biotherapeutics alters dairy cows so that certain immune cells carry human DNA that allows people to make antibodies. That upgrade enables the animals to manufacture large quantities of human antibodies against a pathogen protein injected into them, such as the "spike" surface protein of the new coronavirus. The cows' blood contains polyclonal antibodies, a range of the molecules that recognize several parts of the virus. Such antibodies may remain effective—to treat or prevent infections—even if a virus mutates.

Warp Speed's opaque picks

VACCINE RESEARCH | After *The New York Times* reported last week that the Trump administration had chosen five experimental COVID-19 vaccines to fast-track, top scientists linked with the program told *Science* they had not been consulted in the selection. The White House project, Operation Warp Speed, aims to have enough safe, effective product to vaccinate 300 million Americans by January 2021. "It's just odd. They're asking us for advice, which we're providing, but we have no insight into the decision-making," said Peter Hotez, a vaccine researcher at Baylor College of Medicine, who is on a vaccine committee organized by the National Institutes of Health to help Warp Speed run clinical trials. "We're sort of like two parallel universes."

SCIENCEMAG.ORG/TAGS/CORONAVIRUS
Read additional *Science* coverage of the pandemic.

PHOTO: MORTEZA NIKOUBAZL/SIPA/NEWSCOM

DIVERSITY

Essay prompts mass resignations

Sixteen of the 44 advisers to a leading chemical journal resigned this week to protest an article they said promoted "racist and sexist views." The essay in *Angewandte Chemie*, by Tomáš Hudlický of Brock University in St. Catharines, Canada, argued that "preferential status" given to women and minority job candidates is "counter-productive if it results in discrimination against the most meritorious candidates." The German Chemical Society, which publishes the journal, removed the article from its website after a flurry of criticism on social media, apologized for running it, and suspended two editors. The advisory board members, including Nobel laureates Frances Arnold, William Moerner, and Jack Szostak, said they hoped their departure would provide the journal "an opportunity to reconstitute the Board in a way that reflects our broader community and society."

Fishing restrictions lifted

ENVIRONMENT | President Donald Trump last week opened the only fully protected U.S. marine reserve in the Atlantic Ocean to commercial fishing. In 2016, then-President Barack Obama created the 1.2-million-hectare Northeast Canyons and Seamounts Marine National Monument, phasing out fishing for lobster, crab, and finfish. Although economic data show the closures have had little impact on New England's seafood industry, fishing groups urged Trump to lift the bans. On 5 June he did, brushing aside evidence that fishing gear threatens whales and other wildlife. Brad Sewell, an oceans policy specialist at the Natural Resources Defense Council, called Trump's move "a political stunt. … There will be zero jobs gained."

Human cells model embryos

DEVELOPMENTAL BIOLOGY | Using tiny blobs of human cells, researchers have glimpsed an understudied stage of human development: gastrulation. During its third week of development, an embryo becomes multilayered and elongates along a head-to-tail axis. Genetic abnormalities or chemical exposures can lead to birth defects or pregnancy loss at this "gastrula" stage. Ethical limits prevent researchers from growing human embryos in a dish past 14 days. But with carefully timed chemical cues, researchers have coaxed human embryonic stem cells to form "gastruloids" that mimic a gastrula's elongated shape and signature cell types, they report this week in *Nature*. The work avoids many ethical concerns about lab-grown embryo-like structures, the researchers say, because these gastruloids lack cells that would become a brain or allow an embryo to implant in a uterus.

PHOTO: NAOMI MORRIS

Human embryonic stem cells formed structures that can model early development.

Iceland company ends whaling

CONSERVATION | Whaling in Iceland, one of just three countries that permit commercial killing of large whales, is dwindling. Both of the country's remaining whaling firms say they won't hunt this year, citing Iceland's move to expand coastal whale sanctuaries and poor market conditions for whale meat. One of the companies, IP-Utgerd, announced on 24 April it is permanently ending its hunts, saying the 2017 expansion of Iceland's marine sanctuaries made killing minke whales too expensive. The other firm, Hvalur hf, which kills fin whales and exports the meat to Japan, said it won't set sail for the second year in a row. Hvalur cannot compete with Japanese whaling firms, which are subsidized by Japan's government, Hvalur's CEO, Kristján Loftsson, told the newspaper *Morgunblaðið*. Most other countries that had whaled agreed in 1986 to an international ban. Only Japan, Iceland, and Norway still permit commercial whaling.

THREE QS

#BlackBirdersWeek highlights field scientists' challenges

Last week, black scientists and recreational birders flocked to Twitter for #BlackBirdersWeek. The event came after an incident in New York City's Central Park in which a white woman called the police on a black man who, while birding, had asked her to leash her dog. For black scientists, his experience was familiar, says Corina Newsome, a graduate student at Georgia Southern University, Statesboro, who co-organized the online event. For *Science*'s full interview, see https://scim.ag/37BlackBirders.

Q: How did you come up with #BlackBirdersWeek?

A: The idea came from a group that I'm part of, which includes probably 100 black scientists or outdoor enthusiasts. The incident in Central Park was something that all of us have experienced. When we started using these hashtags like #BlackInNature or #BlackBirdersWeek, it brought people from literally around the world who are black into the conversation.

Q: How can academia better support and recruit black scientists?

A: I've heard of white professors who make a rude, racist, or racially charged comment and no one holds them accountable. It is very important that scientists hold their white colleagues accountable. I also think that it's important for faculty members to advocate to invest into strengthening diversity and equity initiatives.

Q: Has race or racism affected you as a field biologist?

A: Where my field site is, there is a plantation museum. And not one that is trying to condemn the racist history. … There are confederate flags everywhere. What made it even more horrible recently is that Ahmaud Arbery was killed [in February] down the street from my field site. Now every time I drive down there, it's a very disorienting feeling because I think about how my people are dying at the hands of a white supremacy. And I'm out here studying birds, and I feel guilty for even engaging in something as disconnected as studying nest predation in seaside sparrows. [After Arbery's death,] I was sitting outside the marsh weeping and sobbing, feeling like I can't even enjoy this—this thing that I've worked so hard to do.

IN DEPTH

The Germans Trias i Pujol University Hospital near Barcelona, Spain, where a prevention trial with hydroxychloroquine took place.

COVID-19

Big studies dim hopes for hydroxychloroquine

Amid politicization and scandal, a disappointing scientific picture is emerging

By **Kai Kupferschmidt**

Through the fog of alleged misconduct, hope, hype, and politicization that surrounds hydroxychloroquine, the malaria drug touted as a COVID-19 treatment, a scientific picture is now emerging.

Praised by presidents as a potential miracle cure and dismissed by others as a deadly distraction, hydroxychloroquine was spared a seeming death blow last week. On 4 June, after critics challenged the data, *The Lancet* suddenly retracted a paper that had suggested the drug increased the death rate in COVID-19 patients (see p. 1167), a finding that had stopped many clinical trials in their tracks. But now three large studies, two in people exposed to the virus and at risk of infection and the other in severely ill patients, show no benefit from the drug. Coming on top of earlier smaller trials with disappointing findings, the new results mean it's time to move on, some scientists say, and end most of the trials still in progress.

"It just seems like we are ignoring signal after signal," says Eric Topol, director of the Scripps Translational Science Institute. U.S. President Donald Trump's promotion of it led to a scientific "obsession" with hydroxychloroquine despite thin evidence for its promise, he says. "We'd be better off shifting our attention to drugs that might actually work." Peter Kremsner of the University of Tübingen agrees hydroxychloroquine "certainly isn't a wonder drug." The new results left him "wrestling" with the question of whether to proceed with two hydroxychloroquine trials, one in hospitals and the other in patients with milder illness at home.

Science's COVID-19 coverage is supported by the Pulitzer Center.

Hydroxychloroquine and its sister drug chloroquine have been used against malaria and other diseases for decades. The first evidence that they might work against SARS-CoV-2 came from test tube data. Since then, hundreds of trials have been launched around the globe. Scientists are trying the drugs in low doses and high doses; alone or combined with the antibiotic azithromycin, the antiviral compound favipiravir, or other drugs; and in patients with mild or severe disease, health care workers, pregnant women, and people living with HIV.

On 5 June, researchers in the United Kingdom announced the results from the largest trial yet, Recovery, in a press release. In a group of 1542 hospitalized patients treated with hydroxychloroquine, 25.7% had died after 28 days, compared with 23.5% in a group of 3132 patients who had only received standard care. "These data convincingly rule out any meaningful mortality benefit," wrote the investigators, who ended the study early and promised to publish the full results as soon as possible.

The results are persuading some doctors to stop using the drug for COVID-19. "The Recovery trial, in addition to the signals from other studies we have received so far, are enough to convince me to not offer hydroxychloroquine to hospitalized patients," Nahid Bhadelia, a physician at Boston Medical Center, wrote in an email. Martin Landray of the University of Oxford, one of Recovery's principal investigators, agrees: "If you, your spouse, your mother gets admitted to hospital and is offered hydroxychloroquine, don't take it," he says.

But some scientists say they want to see the full data before making up their minds. About one in four patients died in both arms of the study, Kremsner notes—a very high rate, suggesting they were gravely ill when treatment started. Nicholas White of Mahidol University in Bangkok, who also studies hydroxychloroquine, agrees the full data need evaluation. "But overall, it's very

PHOTO: FELIPE DANA/ASSOCIATED PRESS

unlikely, in my view right now sitting here, that anything's going to change," he says.

Another hope for hydroxychloroquine, that it might prevent people exposed to the virus from getting sick, also faded last week when David Boulware of the University of Minnesota, Twin Cities, and colleagues published the results of the largest study to date of this strategy, called post-exposure prophylaxis (PEP). The researchers sent either hydroxychloroquine or a placebo by mail to 821 people who had been in close contact with a COVID-19 patient for more than 10 minutes without proper protection. They reported in *The New England Journal of Medicine* that 12% of the people who took the drug went on to develop COVID-19 symptoms, versus 14% in a placebo group, a difference that was not statistically significant.

A second large PEP trial has come up empty as well, its leader tells *Science*. Carried out in Barcelona, Spain, that study randomized more than 2300 people exposed to the virus to either hydroxychloroquine or the usual care. There was no significant difference between the number of people in each group who developed COVID-19, says Oriol Mitjà of the Germans Trias i Pujol University Hospital. Mitjà says he has submitted the results for publication.

"If you, your spouse, your mother gets admitted to hospital and is offered hydroxychloroquine, don't take it."

Martin Landray, University of Oxford

The data are important because they come from large randomized trials. So far, most data came from small trials or case series. A meta-analysis of 24 such studies published in the *Annals of Internal Medicine* concluded there was "insufficient and often conflicting evidence on the benefits and harms of using hydroxychloroquine or chloroquine to treat COVID-19."

The new findings raise questions about whether to stop other trials. Most are much smaller than Recovery, and thus less powerful; their outcomes are unlikely to change many minds. And continuing the trials may prevent researchers from testing drugs with a better chance of working and robs patients of the chance to try those. Landray says the World Health Organization (WHO) is now likely to end the hydroxychloroquine arm of its large COVID-19 treatment trial, named Solidarity. "I think the decision is pretty obvious," he says. WHO says it is considering the issue.

There is one exception. Many researchers agree that a good case can be made for continuing to test whether hydroxychloroquine can prevent infection if given to people just in case they get exposed to the virus, for instance on the job at a hospital—a strategy called pre-exposure prophylaxis (PrEP). "You have a much better chance of preventing something with a weak drug than you have of curing a fully established infection," says White, who runs one of the largest PrEP trials. He notes that doxycycline, an antibiotic, has long been used in malaria prophylaxis. "We would never treat anybody with it, it's too weak. But it's a very good prophylactic."

Landray, however, is on the fence about continuing prophylaxis trials: "I suspect it's one of these decisions where there isn't a right or wrong." It's an important question, Bhadelia says, because an effective PrEP drug could have a major impact on the pandemic. Hydroxychloroquine, a cheap and widely available drug, is one of the few compounds that could fit the bill.

But the *Lancet* paper, despite its retraction, will make it more difficult to continue current trials, White laments. Published on 22 May, the study claimed, supposedly based on data from 96,000 patients around the world, that hydroxychloroquine and chloroquine, whether given alone or in combination with another drug, caused a steep increase in deaths. That led many regulatory agencies to ask scientists to halt their trials and make sure they were not harming their patients. Recovery and Solidarity were temporarily halted but resumed after a safety committee took a look at the data.

Many other trials are still on pause. U.K. regulators, for instance, have asked for a raft of additional safeguards, says Joseph Cheriyan, a clinical pharmacologist at Cambridge University Hospital and principal investigator of a PrEP trial in health care workers. That study already excluded patients who take any one of several dozens of drugs, but Cheriyan says regulators have asked for more changes, which will set the trial back weeks. And despite the *Lancet* retraction, the alarming headlines about the drug's risks have made it much more difficult to convince people to participate in a trial, White says. "I just think these trials have been really badly damaged and some of them may never restart."

The problem for scientists is that there's such a rush to find treatments for the rapidly spreading virus, Mitjà says: "The pressure is immense." Yet that shouldn't stop researchers from properly analyzing data and making carefully considered decisions, White says. "We don't always have to act today," he says. "Let's not panic." ■

COVID-19

Authors, elite journals under fire after major retractions

Editors, co-authors missed warning signs, critics say

By **Charles Piller** *and* **John Travis**

Last month, Mandeep Mehra, Amit N. Patel, and Sapan Desai were riding high, with shared co-authorships on major new papers in *The Lancet* and *The New England Journal of Medicine* (*NEJM*) and an influential preprint. Drawing on what appeared to be a vast patient data trove from hospitals around the world, the papers delivered seemingly definitive news about whether already approved drugs were safe for COVID-19 patients, or effective against the disease.

Now, the two journal papers have been retracted, the preprint taken down, and Patel's academic affiliation severed. The three physician-scientists are under the microscope as a shocked scientific community evaluates what may be the first major episode of research fraud in the pandemic. The journals are receiving withering criticism for what some call a failure of editorial processes and peer review. The retractions are "unnerving and disturbing," says Leigh Turner, a bioethicist at the University of Minnesota, Twin Cities. The rush to publish on COVID-19 has exposed a lack of rigor that has reached "elite journals at the top of the academic pyramid," he says.

The retracted *NEJM* paper "had external peer review and statistical review, as well as scientific and manuscript editing," an *NEJM* spokesperson says. *The Lancet* did not comment on its review process. Neither journal notes submission or acceptance dates for papers, but a spokesperson for Mehra says reviews for each paper took about 1 month.

By publishing only author retraction statements, the journals "didn't show any self-reflection, any introspection," Turner says. To him, the case also raises a bigger question about how much access to key data each journal should require—and whether all co-authors should have full access to a data set. "The less access they have, the greater the chances that there will be

errors, data fabrication, or outright fraud."

Publication of the *Lancet* paper abruptly halted many trials of hydroxychloroquine, the antimalarial touted by President Donald Trump, because of its finding that COVID-19 patients receiving the drug had a greater death rate than a control group (see p. 1166). The *NEJM* paper exonerated blood pressure drugs that some thought might worsen COVID-19, and the preprint found that mortality was dramatically reduced in COVID-19 patients receiving the parasite drug ivermectin, which drove huge demand for the medicine in Latin America (*Science*, 5 June, p. 1041).

Mehra, Patel, and Desai were the only scientists on more than one of the three papers, and all of the other co-authors are linked to at least one of the trio. After critics discovered anomalies in the data and wondered how Surgisphere, Desai's small company, could have amassed and analyzed tens of thousands of hospital records from around the world, the core authors promised independent data audits. But Surgisphere declined to make the firm's database and hospital agreements available, prompting the journal retractions. "We can no longer vouch for the veracity of the primary data sources," Mehra, Patel, and a third author wrote in the *Lancet* retraction.

The ivermectin study quietly vanished from the SSRN preprint server. "There's no retraction letter. But its ghost lives on in Latin America," says tropical disease physician Carlos Chaccour of the Barcelona Institute for Global Health, who, with colleagues, raised questions about the preprint. (African physicians who developed a COVID-19 severity rating system with Surgisphere's help withdrew the tool last week.)

Desai, Mehra, and Patel had never before published together, and that should have been a red flag to any journal, says Jerome Kassirer, editor-in-chief of *NEJM* during the 1990s. Co-authors of high-profile papers normally share subject area expertise or have clear professional ties, he says, calling the collaboration of the apparently disparate individuals "completely bizarre."

Prior to the retractions, Desai, a science-fiction writer, entrepreneur, and vascular surgeon, had defended Surgisphere and its database. Neither Mehra, a highly respected scientist at Harvard University and Brigham & Women's Hospital, nor Patel, a little known cardiac surgeon who recently resigned from an unpaid adjunct position at the University of Utah, has talked to the press. But Mehra apologized in a statement. "I did not do enough to ensure that the data source was appropriate for this use. For that, and for all the disruptions—both directly and indirectly—I am truly sorry."

As CEO of Surgisphere, Desai has received the most scrutiny. He started the company in 2007 as a medical resident at Duke University. It initially produced medical guides. In 2010, under the firm's auspices, he founded the *Journal of Surgical Radiology*, which folded in 2013. Its articles have been cited only 29 times, according to Scimago, a journal rating service. Yet an undated Surgisphere web page, no longer accessible, said the online-only publication had 50,000 subscribers and nearly 1 million page views monthly.

Surgisphere also claimed to have gathered and analyzed data on nearly 100,000 patients at some 700 hospitals worldwide. But no hospitals have acknowledged giving data to the firm. National Health Service Scotland, noted in a case study on the company's website, tells *Science* that none of its hospitals worked with Surgisphere. It will ask the firm to remove a website image of a Glasgow hospital.

While he was CEO of Surgisphere, Desai practiced at U.S. hospitals; Illinois court records show he is facing two malpractice suits filed last year. He spoke often at medical conferences, impressing more senior researchers. In a conference talk last year on "millennials" in vascular surgery, Gilbert Upchurch, chair of the University of Florida's surgery department, brought up Desai, saying he had never worked with him, but had mentored him remotely and they had an online friendship. Upchurch placed Desai in a group of "amazing and talented young vascular surgeons."

But another physician scientist who worked closely with Desai several years ago says the more time people spent with him, the greater their doubts. "Just about everyone who knew him would say: 'I just didn't have a good feeling about him.' … After they'd been with him, most people dissociated themselves from him," says the person, who declined to be named to avoid personal and institutional embarrassment.

Patel started as a full-time faculty member at the University of Utah in 2008, gained tenure in 2013, and in late 2016 moved to University of Miami's Miller School of Medicine as head of cardiac surgery. Recently, he returned to the University of Utah. Patel is also an unpaid collaborator on a trial using stem cells from umbilical cord blood to treat COVID-19, according to Camillo Ricordi, chief of cellular transplantation at the University of Miami and a principal investigator on that trial. Ricordi praises Patel's work in regenerative medicine.

Much of Patel's decadeslong research career has involved experimental stem cell therapies purported to, for example, treat or cure heart disease and sexual dysfunction, or reverse aging. The therapies were sometimes sold with limited evidence of efficacy. Patel has never received National Institutes of Health funding, according to the agency's database. And of more than 100 publications listed on his University of Utah online profile—which the school removed last week upon his leaving—nearly two-thirds were actually co-authored by other scientists who share Patel's surname, *Science* found.

Patel recently tweeted that he is "related to Dr. Desai by marriage," but added that he remains in the dark about the Surgisphere data. Mehra, author of more than 200 scholarly articles and editor of a transplantation journal, enjoys considerable support even after the retractions. "I've never had any indication whatsoever that he would do anything unethical," says Keith Aaronson, a cardiologist at the University of Michigan, Ann Arbor, who has collaborated with Mehra.

Mehra says he met Patel in "academic and medical circles" and that Patel connected him to Desai. In journal papers, including the retracted ones, Mehra also acknowledged receiving consulting fees from Triple-Gene, a gene therapy company Patel co-founded. "I think [Mehra] just fell into this—perhaps a little naïvely," says another collaborator, surgeon Daniel Goldstein of the Albert Einstein College of Medicine.

But Kassirer faults Mehra for apparently letting ambition get the best of him. "If you're a scientist and you're going to sign on to a project, by God you should know what the data are," Kassirer says. ■

Sapan Desai, Mandeep Mehra, and Amit N. Patel (left to right) co-authored retracted COVID-19 papers.

PHOTOS: (LEFT TO RIGHT) SOUTHERN ILLINOIS UNIVERSITY SCHOOL OF MEDICINE; © STU ROSNER; LAURA SEITZ/*DESERET NEWS*

With reporting by Kelly Servick and Martin Enserink. This story was supported by the *Science* Fund for Investigative Reporting.

Mink populations burgeon in the spring, when pups are born, raising concerns about new SARS-CoV-2 outbreaks.

PHOTO: RUSLAN SHAMUKOV/TASS VIA GETTY IMAGES

COVID-19

Coronavirus rips through Dutch mink farms, triggering culls

Public concerns rise as two workers catch the disease

By **Martin Enserink**, *in Lelystad, the Netherlands*

In a sad sideshow to the COVID-19 pandemic, authorities in the Netherlands began to gas tens of thousands of mink on 6 June, most of them pups born only weeks ago. SARS-CoV-2 has attacked farms that raise the animals for fur, and the Dutch government worries infected mink could become a viral reservoir that could cause new outbreaks in humans.

The mink outbreaks are "spillover" from the human pandemic—a zoonosis in reverse that has offered scientists in the Netherlands a unique chance to study how the virus jumps between species and burns through large animal populations.

But they're also a public health problem. Genetic and epidemiological sleuthing has shown that at least two farm workers have caught the virus from mink—the only patients anywhere known to have become infected by animals. SARS-CoV-2 can infect other animals, including cats, dogs, tigers, hamsters, ferrets, and macaques, but there are no known cases of transmission from these species back into the human population. (The virus originally spread to humans from an as-yet-unidentified animal species.)

The first two mink outbreaks were reported on 23 and 25 April at farms holding 12,000 and 7500 animals, respectively. More mink were dying than usual, and some had nasal discharge or difficulty breathing. In both cases, the virus was introduced by a farm worker who had COVID-19. Today, it has struck 12 of about 130 Dutch mink farms. Once COVID-19 reaches a farm, the virus appears to spread like wildfire, even though the animals are housed in separate cages. Scientists suspect it moves via infectious droplets, on feed or bedding, or in dust containing fecal matter.

That mink are susceptible wasn't a surprise, because they are closely related to ferrets, says Wim van der Poel of Wageningen University & Research, which has an animal health laboratory here. (Both mink and ferrets can also contract human influenza viruses.) Like humans, infected mink can show no symptoms, or develop severe problems, including pneumonia. Mortality was negligible at one farm and almost 10% at another. "That's strange—we don't really understand it," says virologist Marion Koopmans of Erasmus Medical Center in Rotterdam. Feral cats roaming the farms—and stealing the mink's food—were found to be infected as well.

The Netherlands is the only country so far to have reported SARS-CoV-2 in mink. In Denmark, the world's largest mink producer, "We have not recorded any similar disease or outbreaks," says Anne Sofie Hammer, a veterinary scientist at the University of Copenhagen. Neither has China, the second largest producer, says virologist Chen Hualan of the Chinese Academy of Agricultural Sciences. (Hubei, the province hardest hit by COVID-19, does not have mink farms, she notes.)

The Dutch outbreaks are giving scientists a chance to study how the virus adapts as it spreads through a large, dense population. In some other animal viruses, such conditions trigger an evolution toward a more virulent form, because the virus isn't penalized if it kills a host animal quickly as long as it can easily jump to the next one. (Avian influenza, for instance, usually spreads as a mild disease in wild birds but can become highly pathogenic when it lands in a poultry barn.) Although SARS-CoV-2 is undergoing plenty of mutations as it spreads through mink, its virulence shows no signs of increasing.

Even so, the Dutch outbreaks have alarmed people in North Brabant province, where mink farms are concentrated. The region's burgeoning goat industry caused the world's largest human epidemic of Q fever between 2007 and 2009 (*Science*, 15 January 2010, p. 266). Anxious citizens feared a repeat with SARS-CoV-2 and mink. But *Coxiella burnetii*, the bacterium that causes Q fever, forms hardy spores that wafted out of barns and blew off fields fertilized with goat manure. SARS-CoV-2 is far more fragile; environmental sampling has not turned up any virus outside mink sheds, says veterinary epidemiologist Arjan Stegeman of Utrecht University, who leads the research on mink outbreaks. Whereas farm workers should wear protective equipment, the population at large is at very low risk, Stegeman says.

Eventually, the virus seems to burn itself out at every farm, once more than 90% of the animals have contracted it and developed antibodies. Combined with the low mortality rate, that means the outbreaks are far less devastating for farmers than, for instance, bird flu in poultry or foot-and-mouth disease in cattle.

Even though just two of the Netherlands's nearly 50,000 confirmed human COVID-19 cases have been linked to the farms, the government decided to cull the animals because the problem could become bigger in the months ahead. Female mink give birth in April and May, leading to a sixfold increase in populations. Antibodies in their mother's milk probably protect pups for a while, but they might become vulnerable later to any virus lingering at the farm. "That could mean there's a second wave in minks in the fall," Van der Poel says—raising the risk of more human cases. The mink are culled by gassing them with carbon monoxide; the Dutch government will compensate farmers.

In the long run, their businesses were doomed anyway: A law approved by the Dutch parliament in 2012 bans mink farming as of 2024 for ethical reasons. The affected farmers may be allowed to reopen their farms for another 3 years if tests conclusively show the virus is gone—or they can decide to throw in the towel now. ■

COVID-19

Vaccines that use human fetal cells draw fire

Abortion opponents urge United States and Canada to avoid "ethically-tainted" cell lines

By **Meredith Wadman**

Senior Catholic leaders in the United States and Canada, along with other antiabortion groups, are raising ethical objections to promising COVID-19 vaccine candidates that are manufactured using cells derived from human fetuses electively aborted decades ago. They have not sought to block government funding for the vaccines, which include two candidate vaccines that the Trump administration plans to support with an investment of up to $1.7 billion, as well as a third candidate made by a Chinese company in collaboration with Canada's National Research Council (NRC). But they are urging funders and policymakers to ensure that companies develop other vaccines that do not rely on human fetal cell lines and, in the United States, asking the government to "incentivize" firms to make only vaccines that don't rely on fetal cells.

"It is critically important that Americans have access to a vaccine that is produced ethically: no American should be forced to choose between being vaccinated against this potentially deadly virus and violating his or her conscience," the U.S. Conference of Catholic Bishops (USCCB) and 20 other religious, medical, and political organizations wrote to Stephen Hahn, commissioner of the U.S. Food and Drug Administration (FDA), in April. "Thankfully, other [COVID-19] vaccines … utilize cell lines not connected to unethical procedures and methods."

"We urge your government to fund the development of vaccines that do not create an ethical dilemma for many Canadians," wrote Richard Gagnon, archbishop of Winnipeg and president of the Canadian Conference of Catholic Bishops, and 17 other antiabortion religious, medical, and political groups and individuals in a 21 May letter to Prime Minister Justin Trudeau. "The … manufacture of vaccines using such ethically-tainted human cell lines demonstrates profound disrespect for the dignity of the human person."

FDA replied to USCCB on 11 May, writing that, "An inability to use these cells … would deprive the United States of lifesaving vaccines, and … adversely impact" public health. In Canada, the health ministry has promised to respond to the letter to Trudeau, says Moira McQueen, executive director of the Canadian Catholic Bioethics Institute and lead signatory on the letter.

In response to lobbying by antiabortion groups, the Trump administration last year barred U.S. government scientists from using human fetal cell lines from new elective abortions in their work. But the administration has not banned the use of fetal cell lines derived from abortions decades ago. Such cell lines have been used since the 1960s to manufacture vaccines, including current vaccines against rubella, chickenpox, hepatitis A, and shingles.

Now, research groups around the world are working to develop more than 130 candidate vaccines against COVID-19, according to the World Health Organization. At least six of those candidates use one of two human fetal cell lines: HEK-293, a kidney cell line widely used in research and industry that comes from a fetus aborted in about 1972; and PER.C6, a proprietary cell line owned by Janssen, a subsidiary of Johnson & Johnson, developed from retinal cells from an 18-week-old fetus aborted in 1985. Both cell lines were developed in the lab of molecular biologist Alex van der Eb at Leiden University.

Two of the six vaccines have entered human trials (see table, below). Five are made by using human fetal cells as "factories" to make adenoviruses that carry genes from SARS-CoV-2, the virus that causes COVID-19. The adenoviruses, which are disabled so they can't replicate, are given as a vaccine; recipients' cells then produce proteins from the coronavirus, hopefully triggering a protective immune response.

The sixth vaccine, which could enter human trials this summer, is a protein subunit vaccine. Researchers use HEK-293 cells to make pieces of the spike protein that studs the coronavirus' surface. To trigger an immune response, the vaccine is delivered through a skin patch with 400 tiny needles.

Human fetal cells are key to producing both types of vaccines. For the protein subunit vaccine, "Cultured [nonhuman] animal cells can produce the same proteins, but they would be decorated with different sugar molecules, which … runs the risk of failing to evoke a robust and specific immune response," says Andrea Gambotto, a vaccine scientist at the University of Pittsburgh School of Medicine and lead developer of the vaccine. (Of the developers of the six vaccines, only Gambotto responded to a request for comment.)

David Prentice, vice president and research director at the Charlotte Lozier Institute, which opposes abortion, notes researchers making adenovirus vaccines have modified HEK-293 cells to be adept at packaging new genes—such as those that direct cells to assemble the coronavirus spike protein—into adenoviruses. But he adds that other technologies are available, including using fetal cells captured from amniocentesis.

"The use of cells from electively aborted fetuses … makes these … COVID-19 vaccine programs unethical, because they exploit the innocent human beings who were aborted,"

In contention

At least six COVID-19 vaccine candidates use cells from fetuses aborted decades ago.

DEVELOPER	VACCINE TYPE	FETAL CELLS USED	HUMAN TRIALS	POTENTIAL U.S. FUNDING	WARP SPEED PICK
CanSino Biologics, Inc./Beijing Institute of Biotechnology	Replication-deficient adenovirus	HEK-293	Yes (phase II)	No	No
University of Oxford/ AstraZeneca	Replication-deficient adenovirus	HEK-293	Yes (phase II/III)	$1.2 billion	Yes (short list)
Janssen Research & Development USA	Replication-deficient adenovirus	PER.C6	No	$456 million	Yes (short list)
University of Pittsburgh	Protein subunit	HEK-293	No	No	No
ImmunityBio/ NantKwest	Replication-deficient adenovirus	HEK-293 or derivative E.C7	No	No	Yes (long list)
altimmune	Replication-deficient adenovirus	PER.C6	Expected this month	No	No

Prentice and a co-author—molecular biologist James Sherley, a Lozier Institute associate scholar and director of the adult stem cell company Asymmetrex—wrote last month.

But Arthur Caplan, a bioethicist at the New York University School of Medicine, counters: "There are better ways to win the abortion wars than telling people not to use a vaccine. These are long-over abortions. These cells are decades old, and even major religious leaders like the pope have acknowledged that for the greater good it's not worth the symbolism to put the community at risk."

The Vatican's Pontifical Academy for Life declared in 2005 and reaffirmed in 2017 that, in the absence of alternatives, Catholics could in good conscience receive vaccines made using historical human fetal cell lines.

One of the six vaccines, made by the Chinese company CanSino Biologics, was the first COVID-19 vaccine to enter phase II human trials. It uses HEK-293 cells that the company licensed from Canada's NRC, which developed the cells. (The firm has previously used HEK-293 cells from NRC to develop an approved Ebola vaccine.) NRC is now collaborating with CanSino Biologics, preparing to run trials of the vaccine in Canada and scale up production facilities.

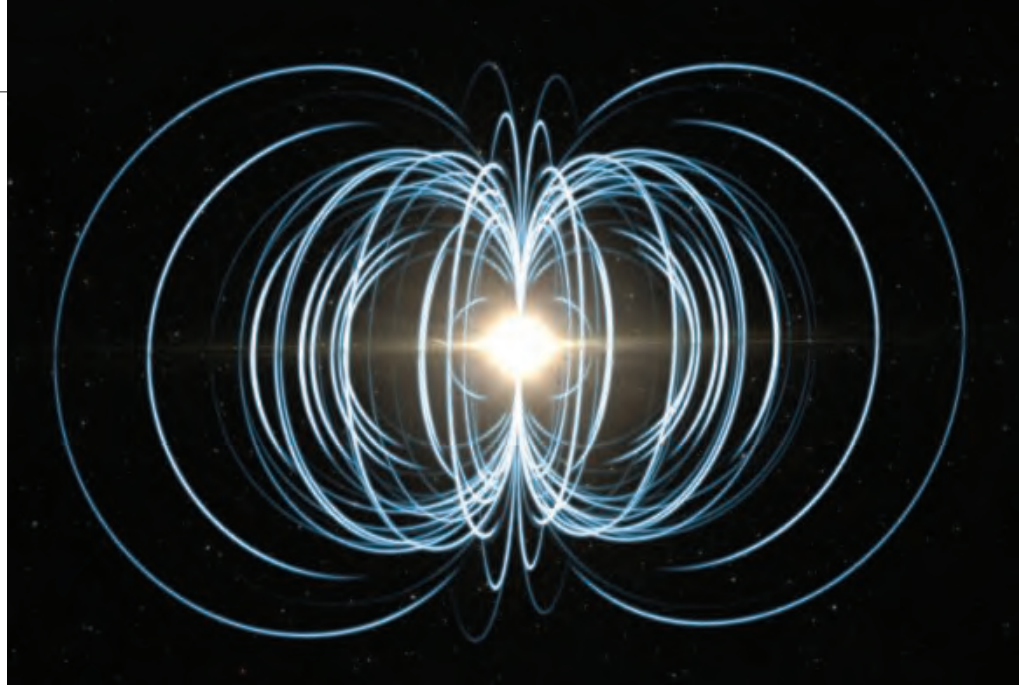
Two vaccines that have drawn criticism from antiabortion groups are on a short list of candidates to get financial and logistical support from the U.S. government under the White House's Operation Warp Speed, which aims to deliver at least one approved COVID-19 vaccine by January 2021, according to a 3 June report in *The New York Times*. One, made by Janssen Research & Development, uses PER.C6 cells. The second, from University of Oxford researchers and AstraZeneca, uses HEK-293 cells. Both have received U.S. government commitments of, respectively, $456 million and $1.2 billion, if they meet milestones, through the Biomedical Advanced Research Development Authority.

Another vaccine that relies on HEK-293 cells made a Warp Speed long list of 14 candidates, according to a press release from NantKwest, one of two companies owned by billionaire scientist Patrick Soon-Shiong that are developing the vaccine.

Prentice believes the government should think twice about supporting such vaccines. "As they are choosing … what vaccines to move ahead, they should at least recognize that there is some portion of the population who would like an alternative vaccine they can take in good conscience," he says.

Caplan disagrees. "If you are going to say the government shouldn't fund things that a minority of people object to, you will have a very long list of things that won't get funded by the government, from research on weapons of war to contraceptive research." ■

IMAGE: © PITRIS/DREAMSTIME.COM



Magnetars are collapsed stars with fields 100 million times stronger than that of any magnet on Earth.

ASTRONOMY

Galactic flash points to source of enigmatic fast radio bursts

CHIME telescope catches burst generated by a highly magnetized neutron star in the Milky Way

By **Daniel Clery**

On 28 April, as Earth's rotation swept a Canadian radio telescope across the sky, it watched for mysterious millisecond flashes called fast radio bursts (FRBs). At 7:34 a.m. local time an enormous one appeared, but awkwardly, in the peripheral vision of the scope. "It was way off the edge of the telescope," says Paul Scholz, an astronomer at the University of Toronto and a member of the Canadian Hydrogen Intensity Mapping Experiment (CHIME). Because of its brightness, the team knew its source was nearby. All other FRBs seen so far have erupted in distant galaxies—too far and too fast to figure out what produced them.

The team had a hunch about this one. In previous days, orbiting telescopes had witnessed a Milky Way magnetar—a neutron star with a powerful magnetic field—flinging out bursts of x-rays and gamma rays. The turmoil suggested it might be pulsing with radio waves, too. After some extra data processing, the team determined the FRB was "definitely colocated" with the magnetar, Scholz says. "We were really excited."

The find, announced in a paper posted to the arXiv preprint server on 20 May, could be the missing link in a problem that has puzzled astronomers for more than a decade. It's only a single event and many questions remain, including why this burst was 30 times less energetic than the weakest FRB traced to another galaxy. Yet astronomers are increasingly confident that some, if not all, of these laserlike radio flashes originate from magnetars, collapsed stars with magnetic fields 100 million times stronger than any magnet made on Earth. A magnetar origin would rule out more exotic sources such as supermassive black holes and merging neutron stars. "The game of alternative theories is becoming more and more difficult," says theorist Maxim Lyutikov of Purdue University. "For the majority it's a decided question: It's magnetars."

The first FRB was detected in 2007, and astronomers have tallied a little over 100 since then. Their brevity makes them hard to study or trace to a particular celestial object. But several FRBs have been found to repeat, giving astronomers a chance to identify their host galaxy. And in the past year or two, wide-field telescopes such as CHIME, designed to survey large swaths of the sky, have begun to boost the number of detections substantially (*Science*, 15 March 2019, p. 1138).

Another one of these wide-field telescope is the Survey for Transient Astronomical Radio Emission 2—three radio antennas scattered across the western United States. It also spotted the 28 April burst and measured its energy, according to a second preprint, posted to arXiv on 21 May. Deepening the mystery, the proposed source of the burst, a magnetar called SGR 1935+2154, is "nothing special," Lyutikov says. Only five of the 30 known magnetars in the Milky Way have been seen to emit weak radio signals, and SGR 1935+2154 is not one of them.

Yet theorists are already jostling to explain how low-key magnetars could power the monumental blasts. "Pretty much every modeler who had previously considered how magnetars could generate an FRB has now said, within days, that they are right," says Victoria Kaspi, an astronomer at McGill University and a CHIME leader. "They can't all be right."

One model, proposed by Brian Metzger of Columbia University and his colleagues, suggests that magnetars sometimes emit bursts of particles near the speed of light, akin to the great puffs of plasma the Sun belches in coronal mass ejections. When a burst slams into a shell of material emitted earlier, it creates a shock wave that causes electrons to spiral around magnetic field lines, generating a powerful laserlike radio pulse. Metzger's group had not applied its model to something as weak as SGR 1935+2154, but when it did, "it worked OK," he says. The model, Metzger says, can also explain why the magnetar's x-ray pulse was 100,000 times more energetic than its radio one.

Lyutikov believes the action occurs much closer to the magnetar's surface. In 2002—years before the discovery of the first FRBs—he proposed an engine based on magnetic reconnection, in which field lines break and reconnect in new configurations. On the surface of the Sun, the phenomenon drives powerful flares. On a neutron star, Lyutikov says, it could generate the nearly simultaneous bursts of both x-rays and radio waves seen from SGR 1935+2154, although it does not yet account for the laserlike beaming.

Observers are now training their telescopes toward nearby galaxies, hoping to catch more flyweight FRBs. Although the 28 April flash was puny compared with most extragalactic FRBs, it does at least show that magnetars are capable of something more dramatic than their usual feeble radio signals. "This goes a long way to bridging that gap," says Jason Hessels of the University of Amsterdam. "It all seems a lot more plausible now." ■

A former student alleges her access to Incan knotted textiles required her to please Harvard University's Gary Urton.

PHOTO: © SAM OGDEN

SCIENTIFIC COMMUNITY

Prominent Harvard anthropologist put on leave

Three former students accuse Inca scholar Gary Urton, alleging propositions or "traumatic" affairs

By **Ann Gibbons**

On the morning of 30 May, anthropologist Jade Guedes read an article in the online issue of *The Harvard Crimson* in which a former student alleged that anthropologist Gary Urton had propositioned her. The story triggered Guedes's memory of an incident in July 2012: She had been a Ph.D. student at Harvard University, and Urton, who was soon to be department chair, had invited her to a "tête-à-tête" to discuss her research. Guedes, then 32, had felt "really great—this big professor is interested in my work."

Then she got an email from Urton: "I wonder if you would be interested in something more intimate? … what if I got a hotel room and then we got a bottle of wine and spent an afternoon in conversation and exploration? … I do hope this is not shocking to you, or disturbing."

Guedes panicked, consulted friends and a mental health adviser, and turned Urton down. She did not report his solicitation, but now, reading an account that in some ways paralleled her own, she was filled with rage. A tenured anthropologist at the University of California, San Diego, and the Scripps Institution of Oceanography, she posted Urton's 2012 email on 1 June on Twitter. The next day, she filed a complaint of sexual harassment against Urton with Harvard's Title IX office.

Urton, now 71, an anthropological archaeologist and scholar of Andean culture, said in a statement to *Science*: "As much as I would love to respond to the false allegations that are circulating and destroying my professional reputation, I have been advised not to do so at this time, other than to say that I am truly saddened by the allegations and hope that some day I will have the chance to clear my name." *The Harvard Crimson* has reported that Urton apologized for his behavior to Guedes in a statement to the paper on 2 June.

Guedes wasn't the only one flooded with unwelcome memories last week. "I spent the day shaking; I haven't been able to sleep," says anthropologist Carrie Brezine, a data analyst at the University of Michigan, Ann Arbor, and a former Ph.D. student of Urton's. "I'm so sick this happened to other people."

In conversations with *Science*, Brezine alleged that Urton seduced her at a remote field site in Peru in 2003, when she was 32. She says Urton had hired her to create a database of examples of khipu, complex knotted textiles that the Inca used as a numerical recording system. Brezine says she and Urton had an affair from 2003 to 2009 while he was her employer and later her Ph.D. adviser. She alleges that her work and research on khipu depended on Urton's good will, which was conditional on sex. She is contemplating filing a complaint with Harvard's Title IX office.

The *Crimson* article, by rising junior James Bikales, has triggered an earthquake

in Harvard's anthropology department, long dominated by a small group of senior men. And it has begun to topple Urton, who "until last week, was one of the most respected researchers in the world," says archaeologist Luis Jaime Castillo Butters at the Pontifical Catholic University of Peru.

Urton, a world authority on khipu, has received a MacArthur "genius" grant and a Guggenheim fellowship. He and Castillo Butters, also a former minster of culture in Peru, worked together at the San José de Moro field school in Peru in recent years. Castillo Butters says that "in all fieldwork trips there has been no incident, no complaints."

On 2 June, Harvard Faculty of Arts and Sciences Dean Claudine Gay put Urton on paid administrative leave pending a full review of the allegations. In a statement on 4 June, Gay said Harvard's policy is to not comment on Title IX investigations. She added that "Harvard is deeply committed to providing an environment for all members of our community that is free from sexual harassment and misconduct. … We encourage any member of our community who has experienced inappropriate behavior to come forward."

Also on 2 June, the anthropology department removed Urton as director of undergraduate studies, a post he had held for 2 years after being department chair from 2012 to 2018. On 4 June, 25 faculty members in the department sent Urton a letter calling for him to resign. "The strong evidence put forth in these allegations has destroyed our confidence in your ability to be a teacher, colleague, and productive member of the department," the letter said.

As the *Crimson* article noted, questions about Urton's conduct had emerged in a 2015 lawsuit filed by medical anthropologist Kimberly Theidon, now at the Fletcher School of Tufts University, Medford. She alleged she was denied tenure in 2013 because of gender discrimination and in retaliation for her advocacy for students alleging sexual harassment. In January, the U.S. Court of Appeals for the First Circuit in Boston ruled against her.

However, as evidence of harassment in the department, Theidon's lawsuit brought to light the account that infuriated Guedes. In an affidavit filed in 2016 as a later addition to the lawsuit, an anonymous Harvard Extension School student wrote that in 2011, she asked Urton for a letter of recommendation before she received her final grade in his class. In response, Urton suggested she join him in a private meeting to discuss her "promise" in anthropology. According to the affidavit, Urton instructed the Extension student to go straight to a room he had booked at the Sheraton Commander Hotel near Harvard on 29 December 2011. "I drank the wine he brought and became intoxicated. He in fact made unwelcome sexual advances and I submitted to his advances," she wrote.

She then began a "consensual affair" with Urton that continued off and on until she moved out of the state in 2013. "I would have completed my education at Harvard except that I met Gary Urton," she wrote in the affidavit.

In an interview this week, the Extension student, now 52, says she moved out of the state in 2013 because the affair with Urton was "very traumatic." She says she became severely depressed, was hospitalized, and became suicidal. She quit her job at Harvard Medical School and stopped taking classes. "I knew I would not survive if I didn't leave [Harvard]," she told *Science*.

Her affidavit says that in 2016 she contacted the office of then–Harvard President Drew Gilpin Faust to tell her about Urton "preying on me as a graduate student. … I wanted to be on record to prevent him hurting other people." She did not file a complaint but later provided the affidavit in support of Theidon's case. She told *Science* she believed Theidon's advocacy for women had influenced how Urton, then department chair, had handled Theidon's tenure case.

The alleged interactions with the Extension student happened the year Brezine got her Ph.D. and moved away. Brezine, a weaver, had majored in theoretical math at Reed College. Urton hired her in 2002 as a research assistant to create an important online database on khipu. Brezine traveled with Urton to a remote field site in Peru, where, she alleges, they began an affair. "I thought Gary thought I was special, that I had good ideas, that I could make a contribution to the field," Brezine says.

With the backing of Urton's letter of recommendation as well as a *Science* paper on khipu co-authored with him (*Science*, 12 August 2005, p. 1065), she entered Harvard's Ph.D. program in anthropology. She continued to work for Urton and hoped to do her thesis on khipu. But she says she changed her project because Urton controlled access to the khipu in Harvard's collection as well as the database, a key resource. She feared he would withhold access if he got angry at her.

Kimberly Theidon (top) supported students alleging harassment; Jade Guedes (bottom) posted on Twitter an invitation from Gary Urton to meet in a hotel room.

"If you knew my passion for Inka khipu and wondered why it wasn't in my dissertation, the condition of access to the khipu database was sex," Brezine wrote in an email to *Science*. "Gary made it clear that he could and would revoke my access at any time if I did not perform adequately."

After "years of sexual requests" from Urton, Brezine eventually consulted Harvard's Dean for Graduate Student Affairs Garth McCavana. She recalls that McCavana emphasized that formal complaints rarely worked out in the victim's favor. Complicating her case, Harvard did not prohibit faculty from having sex with students or employees until 2015. (Many other universities banned such relationships years earlier.)

Brezine says she appreciated McCavana's honesty, and decided to do her dissertation on colonial textiles—materials she could access from other researchers—instead. Urton remained her adviser. She left Harvard in 2011 for a postdoc in anthropology at Michigan. "It's difficult for me to overestimate the costs" of her affair with Urton, she says.

McCavana did not respond to requests for comment.

Brezine has decided to speak up and be named because "I am so upset with myself that I didn't pursue it because other people might not have had to go through it. I just didn't know that this was his habit. I thought I was unique."

Theidon argues that Harvard should have followed up on any complaints about Urton and others, even if they weren't official. "I just can't bear it," she says. "How many more students do you think may have been harassed? They are the missing women in the field."

Today, three of the 12 tenured anthropologists in the department, or 25%, are women. The American Anthropological Association reports that 45% of tenured anthropologists are women. Last week, the Harvard department's new leadership convened a committee to investigate how its culture may be problematic for women and to propose recommendations. "We are committed to repairing our relationships and fostering a more supportive, safe, and equitable departmental culture," the department's website says. ■

With reporting by Rodrigo Pérez Ortega.

PHOTOS: (TOP TO BOTTOM) KIMBERLY THEIDON; CHUNBAI ZHANG

COMPUTING CANCER'S WEAK SPOTS

An algorithm to unmask tumors' molecular linchpins is tested in patients *By* **Roxanne Khamsi**

IMAGE: SUMAZIN *ET AL.*, *CELL*, **147**, (2011)

In 2016, doctors invited Eileen Kapotes to join a clinical trial for a drug that had never been used for her disease. Kapotes, a first grade teacher in her 50s, was fighting an aggressive breast cancer that had spread through her body. She had endured grueling treatments over the previous 4 years, including whole-brain radiation therapy. She had also been taking the breast cancer medication Herceptin, but her tumors were still growing. Now, she had a chance to try something radically different: a drug called ruxolitinib, originally designed to treat cancers affecting the blood and bone marrow.

Kapotes's oncologist, Amy Tiersten at Mount Sinai Hospital, was stunned by how well her patient responded to the new drug. It kept her cancer at bay and she had almost no side effects. "I was amazed," Tiersten says.

The ruxolitinib trial was the product of a decadelong quest by Andrea Califano, a systems biologist at Columbia University. Using sophisticated computing, he models the molecular networks that sustain cancer cells and pinpoints proteins called transcription factors that act as linchpins, controlling the behavior of many genes inside a cell. Califano collaborated with cell biologist José Silva, then also at Columbia, to analyze breast cancer samples in a repository of tissues from other patients who had become resistant to Herceptin. Findings of the analysis suggested a transcription factor called STAT3 plays a critical role in those cancers. And ruxolitinib was known to inhibit STAT3.

Other researchers have focused on identifying genetic mutations that drive the disease in an individual patient. Doing so, the thinking goes, can help identify the best drug for each patient. But because of the diversity of cancer-causing mutations across the population, an arsenal of tens of thousands of drugs might be needed to treat everyone.

Califano's approach, by contrast, is a twist on that idea. He has focused instead on identifying a few transcription factors that act as bottlenecks (see graphic, p. 1176). Target those master regulators, as Califano calls them, and you will stop cancer in its tracks, no matter what mutation initially caused it. Oncologists would still need to analyze each patient's mutations to figure out which regulators are at play in their particular cancer, but instead of tens of thousands of drugs, Califano says, they may only need dozens. It's a depersonalized approach to personalized medicine.

The strategy builds on Califano's computational training as a physicist. "We've built algorithms that can reverse engineer the logic of each different tumor so that we know the targets" for drugs, he says. His algorithms are a prime example of systems biology—which uses complicated math to model intricate biological systems, such as gene interactions. It's a field that has generated tremendous interest, but little real-world clinical success.

In 2015, Califano co-founded a company called DarwinHealth that uses his algorithms to guide doctors by identifying the key transcription factors in a patient's tumor and suggesting drugs to target them. His work has earned praise from other researchers, although some note the approach is only in early stages of human testing, and its clinical usefulness remains uncertain. Ed Liu, president and CEO of the Jackson Laboratory, a nonprofit biomedical institute Califano has collaborated with, is optimistic the method will ultimately pay off. "As we develop more and more precise ways to attack those nodes, then the more useful his algorithms will be."

> **"You have more potential combinations of cancerous mutations than atoms in the universe."**
>
> **Andrea Califano**, Columbia University

Califano's approach is about to get its largest test yet. Columbia has allocated $15 million for a trial of 3000 cancer patients within its hospitals over the next 3 years, using DarwinHealth algorithms to analyze each patient's cancer and recommend treatments. "This is probably one of the most exciting moments in my research," Califano says, "because finally we're able to apply this methodology on a scale that is large enough to be able to really learn something in terms of the response of the patient."

IN FALL OF 1958, a young scientist named François Jacob went to his colleague Jacques Monod at the Pasteur Institute in Paris with a hypothesis about how genetic mechanisms might control cell behavior. Both men had renegade tendencies: Jacob had fought Nazis—and been injured—on behalf of the exiled French government in World War II, and Monod, an accomplished rock climber, had taken part in the guerilla activities of the French Resistance. Over the next few years, the pair worked together, and they were the first to demonstrate the idea of genetic circuits. The work ultimately won them a share of the 1965 Nobel Prize in Physiology or Medicine.

In experiments with *Escherichia coli*, Jacob and Monod showed that the gene networks in those bacteria can alter the production of certain enzymes depending on the type of food available. When the sugar lactose was abundant, the bacteria turned on genes that code for the enzymes to metabolize it. But with access only to glucose, a different sugar, the microbes shut down those genes. It was a pioneering demonstration that the activity of individual genes could be either boosted or repressed.

Experiments in later decades helped explain how the cell machinery exerts that control. One key player is transcription factors, proteins that boost or inhibit the activity of other genes. The gene-regulating network of a single cell is far more elaborate than Jacob and Monod had the tools to uncover. The human genome contains 20,000 genes, and an estimated 1500 of those produce transcription factors. That system creates a complex web of on and off switches.

Califano thought that if he could identify the key switches in cancer, he might be able to shut down the catastrophic genetic changes that drive its growth. But after he finished his training as a physicist in 1986, IBM recruited him to spearhead projects in computer vision and artificial intelligence. The building codes at the IBM facility prevented Califano from having an experimental lab to pursue his interests in biology. He left in 2000 and landed at Columbia in 2003. He started to write code to solve the riddle of cancer on the day he arrived.

Nowadays, the data underlying his algorithms come from a method called RNA sequencing (RNA-seq). The method gauges gene activity within cells by sequencing RNA molecules, which act as a proxy for which genes are turned on and off. Algorithms crunch the massive amount of RNA-seq data to reveal which genes are overactive or underactive in cancer compared with healthy tissue. The algorithms then use complex equations to infer patterns of gene interactions and zero in on the transcription factors with the largest influence.

The search for key drivers of cancer isn't easy. Consider a 2018 analysis of more than 9000 samples that reported almost 1.5 million mutations. Genes influence one another in intricate webs and feedback

A visualization of 7000 interacting genes in a brain cancer patient shows the handful with the largest number of interactions (red)—which makes them tempting targets for therapies.

loops, so the number of ways those genetic perturbations might interact in a tumor is vast. "There's, say, 1000 genes that are recurrently mutated across all tumors that may drive cancer, so you have more potential combinations of cancerous mutations than atoms in the universe," Califano says.

The pharmaceutical industry can't make a new drug for each of those unique combinations. (By comparison, 126 new cancer drugs received approval from the Food and Drug Administration [FDA] between 1980 and 2018.) That's why pinpointing the master regulators that are common culprits across cancers is so crucial, Califano says.

At Columbia, he has worked with his former postdoctoral researcher Mariano Alvarez to develop more efficient algorithms for sorting through those networks. The current one, called VIPER—short for virtual inference of protein activity by enriched regulon analysis—has been used in dozens of studies of how vast, interconnected genetic networks have gone awry in bladder, prostate, and lung cancers.

Califano and colleagues recently used the VIPER algorithm to look across RNA-seq data from more than 10,000 individual tumor samples in the Cancer Genome Atlas, a database sponsored by the U.S. government. The team found that different types of cancer have more in common than previously thought. The analysis, now under review for publication, identified 407 transcription factor genes that acted as suspected linchpins across all the cancer samples. Only 20 to 25 of them were implicated in any given cancer—and Califano says fighting the cancer might not require knocking out all those transcription factors: Toppling just a few nodes might be enough.

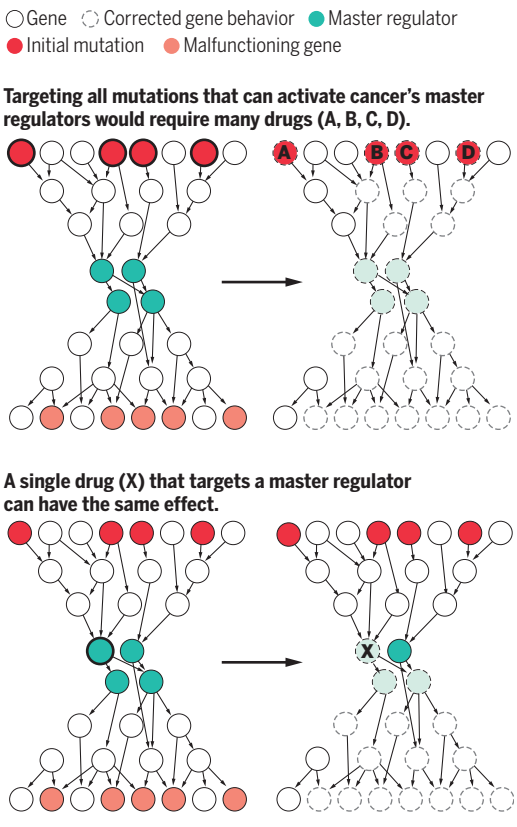
Califano "was among the first to put the complex algorithms out there, and then others have followed," Liu says. A strength of Califano's algorithms is that they look at an entire network of gene products, including RNA and proteins, adds David Tuveson, director of the Cold Spring Harbor Laboratory Cancer Center. Tuveson uses VIPER in his own search for treatments for pancreatic cancer.

Califano, too, hopes to put his algorithms to work for patients. The idea to commercialize that approach began in 2013, on a beach in the British Virgin Islands. There he met a fellow vacationer, Gideon Bosker, a physician who had gotten his start in emergency medicine and later launched a successful medical education company. The pair hit it off, and 2 years later they decided to form DarwinHealth. Columbia licensed the VIPER technology to DarwinHealth, and Bosker put $1.4 million of his money into the venture to get it off the ground. Since then, industry collaborators have sponsored more than a dozen research projects with the algorithms from the company.

DarwinHealth combines Califano's algorithms with a database of information from experiments about how drugs affect multiple genes, compiled through the company's review of the literature and other sources. As long as they follow established rules for patient tissue transfer, doctors around the world can now send a tissue sample to Columbia's pathology department, where RNA is extracted from cells. For $1600, the company generates an "OncoTarget" readout of the individual master regulator that seems to have the biggest sway in the patient's cancer, as well as a more sophisticated "OncoTreat" readout of existing drugs that tamp down the tumor's 25 most activated transcription factors and boost the 25 that are most turned down. The products launched in 2018.

Hitting cancer's choke points

Shutting down a cancer cell's malfunctioning gene network is a tall order if you target mutations too far upstream in the network. But disabling a transcription factor that acts as a master regulator in the cell's genetic circuitry can cause its demise even with just one drug.



GORDON MILLS, who serves as director of precision oncology for the Knight Cancer Institute at Oregon Health & Science University and helped pioneer the field of systems biology, notes that Califano's cancer-fighting algorithms still have to overcome a lot of obstacles. "There's skepticism that we're far enough along to be able to predict the complexity of human disease," says Mills, who has applied Califano's algorithms to data in his own research. "There have been hundreds of algorithms that have failed to truly capture the complexity and heterogeneity of cancer and have not panned out in attempts to move them through to the clinic."

But Califano sees a sign that his search for cancer linchpins will pay off in an unlikely cancer success story: thalidomide. In the mid-1950s, the drug arrived on the market as a sedative to help with sleep and anxiety. Doctors also prescribed it to pregnant women for nausea—which proved devastating because it caused massive birth defects, including missing limbs and heart problems. But thalidomide has made a comeback as a medicine for diseases such as leprosy. In 1997, doctors began to test the drug against multiple myeloma, a cancer affecting white blood cells.

Scientists have since learned more about how thalidomide works. In 2018, they found it prompts a protein complex called cereblon inside cells to mark certain transcription factors for disposal. In multiple myeloma, those transcription factors, IKZF1 and IKZF3, act as linchpins in the genetic network that allows the cancer to thrive. To Califano, thalidomide's success shows the value of finding existing drugs that can target cancer's master regulators.

Candidates are scarce. Whereas many drugs go after proteins that act as enzymes and have easy-to-find active sites to target, transcription factors lack such readily targetable spots, and many researchers have considered them undruggable.

But Califano's Columbia lab is trying to add to the list of potential drugs. Hulking machines with robotic arms process tumor cell samples to do high-throughput screening that looks at how candidate drugs alter the cells' RNA-seq profile and whether the drugs reverse the activity of master regulators. A $12 million supercomputer in the basement of the building, a shared resource for university researchers, analyzes the data.

Bristol Myers Squibb, which manufactures the FDA-approved version of thalidomide for multiple myeloma, has joined the

CREDITS: (GRAPHIC) V. ALTOUNIAN/SCIENCE; (DATA) CALIFANO *ET AL.*, *NATURE REVIEWS CANCER* **17**, 116 (2017)

hunt. It has contracted DarwinHealth to systematically search the pharmaceutical giant's library of compounds for other compounds that might target master regulators.

Additional support for the DarwinHealth approach comes from a recent study by Samir Parekh at the Icahn School of Medicine at Mount Sinai and a team of international collaborators, who recently completed a clinical trial to test a combination of two drugs, dexamethasone and selinexor, for multiple myeloma. The combination only worked in about one-quarter of patients, reducing levels of a myeloma protein in their blood. In a retrospective analysis, the DarwinHealth tools predicted which patients would respond. By assessing RNA-seq data from 12 patients, the tools identified four of the five patients who benefited from the drugs and six of the seven who did not, the researchers reported last year in *The New England Journal of Medicine*.

Morgan Craig, who uses computational approaches to identify new drugs at the University of Montreal, says efforts to understand molecular networks in cancer have the potential to improve personalized medicine. Algorithms like those used by DarwinHealth "may not take over clinical approaches right away," Craig explains. "But it's definitely a step toward trying to do this target identification in a more systematic way."

DARWINHEALTH DOESN'T RUN clinical trials, but for the past 3 years, Califano's lab has tested the company's algorithms in experiments at Columbia. The researchers analyzed RNA-seq data from biopsy samples from more than 100 cancer patients to identify master regulators and suggest drugs that might not normally be considered (much as DarwinHealth's commercial service does). In a few dozen cases, the researchers later tested the drug in mice with a grafted version of a patient's tumor to confirm the drug affected the master regulators as predicted. For five of those patients, doctors felt bold enough to try the algorithm's suggested drug. Each patient had late-stage cancer that had already stopped responding to available treatments.

Four of those five patients responded to the drugs given, at least for a time. For one patient with a meningioma, a tumor that can exert fatal pressure on the brain, the algorithm pointed to etoposide—a drug originally designed to treat lung and ovarian cancer. His tumor stopped growing for more than a year; it then started to slightly rebound and he was put in a different clinical trial. After that, his tumor started to rapidly grow again.

Califano hopes to build on those anecdotal results with the formal clinical trial, now underway at Columbia. The Oncotarget and Oncotreat tests from DarwinHealth will be used with 3000 patients in the Columbia system. Ultimately, the drugs they receive will be chosen by a board of doctors on the basis of readouts either from mutations detected by traditional sequencing or from the VIPER-based algorithm. DarwinHealth will receive no money for the tests to avoid any conflict of interest, given that Califano is part of the company's leadership while working at the university.

Califano and Bosker are also licensing the DarwinHealth tools to other researchers around the world to test against cancer. In January, Beijing Cancer Hospital confirmed it would be using DarwinHealth's tools to guide treatment for patients in clinical trials there. Xiaotian Zhang, an oncologist leading the new study, says that if early results look promising, the research will be expanded to other hospitals. "These clinical studies will all focus on gastrointestinal tumors, particularly gastric and esophageal cancer," Zhang says.

As the DarwinHealth approach goes into more clinical testing sites, more patients like Kapotes will receive drugs never intended to fight their particular cancers. For some people, like her, it might buy precious time. For more than 2 years after she enrolled in the ruxolitinib trial, Kapotes's cancer remained stable. When scans eventually showed her tumor had started to grow again, Tiersten switched her to another medication, which had just received FDA approval. These days, Kapotes is taking time to enjoy her retirement and her family. The newly approved drug she now takes works through a different mechanism, but Kapotes never would have had the chance to take it if not for ruxolitinib. "She hung on long enough because she was in the trial," Tiersten says. ■

Andrea Califano (right) and Gideon Bosker (left) founded DarwinHealth to apply systems biology to cancer.

PHOTO: DARWINHEALTH

Roxanne Khamsi is a science journalist based in Montreal.

INSIGHTS

PERSPECTIVES

PHOTO: RALPH LEE HOPKINS/NATIONAL GEOGRAPHIC IMAGE COLLECTION

ECOLOGY

Costs and benefits of living with predators

Integrating ecology and economics reveals a net monetary benefit from sea otters

By **James A. Estes**[1] *and* **Lilian P. Carswell**[1,2]

Although money isn't everything, today's world operates by means of capitalism. This fundamental truth has led to an intersection between economics and ecology wherein natural resources and ecological processes are often valued in monetary terms. A central challenge to the intersection of ecology and natural resource management is thus twofold—understanding ecological processes and valuing the outcomes. Nowhere has this challenge caused more confusion and debate than in the arena of decision-making about how to conserve and manage large predators. On page 1243 of this issue, Gregr *et al.* (*1*) evaluate the recovery of an apex predator—sea otters in the Canadian Pacific coast—and report that certain economies have been either negatively or positively affected by sea otter restoration, but the overall economic outcomes are positive.

Mother and pup sea otters, *Enhydra lutris*, rest on a bed of kelp onshore in Kachemak Bay, Alaska.

Large apex predators occur in virtually all ecosystems, or once did. They were disproportionately lost or depleted during the age of exploration, discovery, and industrialization. While they are continuing to decline in some cases, they are recovering in many parts of the world following protection and purposeful repatriation. These predators commonly initiate a diverse array of effects on other species and ecological processes (*2*).

Although ecology still has much to learn about the myriad influences of apex predators, most occur through the depletion and fear-induced behavior of their prey. The most well-known ecological pathway from predator through ecosystems is the so-called "trophic cascade," wherein a predator limits the abundance and distribution of its herbivore prey, in turn enhancing the abundance and distribution of plants (*3*). Inasmuch as plants are the foundation for nearly all life on Earth, this link from predator to herbivore to plant extends broadly through nature. Many of the details are reasonably well known for various predators and ecosystems (*4*). Rarely, however, have these processes been rigorously valued on the basis of a common currency. The resulting "apples and oranges" have made it impossible to assess the net costs and benefits of predators to human societies, thereby resulting in irresolvable scientific arguments, irreconcilable management conflicts, and an inability of the responsible management agencies to establish defensible policies. The study of Gregr *et al*. is transformational because it provides a way through this conundrum. The authors conduct a cost-benefit analysis for the extensively studied and well-known case of sea otters and kelp forest ecosystems that accounts not only for shellfish depredation but also for a suite of knock-on effects of this interaction.

Sea otters prey on, and thus limit the size and abundance of, various shellfish, thereby imposing a substantial cost on shellfisheries. However, sea otters also create economic benefits by way of wildlife viewing opportunities and tourism, and through the enhancement of kelp forests that results from a reduction in the number of herbivorous sea urchins. Kelp increases coastal production and provides habitat for other species, thereby enhancing coastal fin fisheries. Kelp also sequesters carbon, thus reducing atmospheric carbon dioxide (CO_2) and ocean acidification. There are numerous other ecological influences of sea otters, although Gregr *et al*. focus on just these, from which they find that the monetary benefits of sea otters outweighed the costs sevenfold. The responsible management agencies in southeast Alaska and British Columbia heretofore have been unable to resolve the severe local conflicts that are developing because of sea otter recovery. Gregr *et al*. do not offer a solution, although they do provide the template for a more holistic and objective means of policy development.

The importance of the analysis of Gregr *et al*. is not so much in the exact findings but as a method for looking beyond depredation in assessing the costs and benefits of living with predators. In the case of sea otters and kelp forests, the web of interactions that emanate from otter predation on shellfish and the otter-urchin-kelp trophic cascade might be expanded to include other interactions that bear on human financial welfare. For example, Pacific herring spawn on kelp, are targeted by fisheries, and provide the nutritional base for other iconic and valued species, such as salmon and whales. The ecological data are not yet sufficient to assess the magnitudes of potential influence by sea otters on these and other species and ecological processes, but in the future they may be. Moreover, the details of Gregr *et al*.'s analysis are specific to the west coast of Vancouver Island, and

[1]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95060, USA. [2]U.S. Fish and Wildlife Service, Ventura, CA 93003, USA. Email: jestes@ucsc.edu

these will surely vary in different places and at different times, perhaps substantially. For example, although tourism is important to many coastal economies along the west coast of Canada and the United States, this isn't so everywhere, nor is it likely that all local communities wish it to be. The ledger of net monetary benefit from sea otters would not switch to a net cost with the removal of tourism, but it would be greatly reduced, given the factors that Gregr *et al.* considered. The important point is that quantifying indirect effects broadens the base of stakeholders in natural resource management by identifying interests that could otherwise be overlooked or dismissed as trivial. Such an expanded perspective alerts the public that they may indeed have a stake in decisions made by natural resource agencies, and it directs policymakers and managers to constituencies they may not have previously consulted.

There is also the larger question of how the monetary costs and benefits of predators affect other ecosystems. This is an important future arena for science and policy. One such scenario involves the gray wolf, which has been the focus of similar controversy and societal discord. The ecological influences of wolves in North American boreal forests are broadly similar to those of sea otters and kelp forests (*5*). One especially topical segment of the web of interactions that emerges from wolves as predators is that of disease. The removal of wolves from the midwestern and eastern United States was likely instrumental in the spread of coyotes into these regions, which in turn depleted fox populations. This led to an increase in small-mammal populations, an associated increase in tick abundance, and the rise of Lyme disease risk (*6*). Such a possibility creates an entirely new dimension to the costs and benefits of living with wolves. The analysis by Gregr *et al.* should spark a new era of ecological-economic research that can be used by natural resource policy-makers and managers to make and defend more rational, equitable, and far-sighted decisions affecting predators. ■

REFERENCES AND NOTES

1. E. J. Gregr *et al.*, *Science* **368**, 1243 (2020).
2. J. A. Estes *et al.*, *Science* **333**, 301 (2011).
3. W. J. Ripple *et al.*, *Trends Ecol. Evol.* **31**, 842 (2016).
4. J. Terborgh, J. A. Estes, Eds., *Trophic Cascades: Predators, Prey, and the Changing Dynamics of Nature* (Island Press, 2010).
5. W. J. Ripple *et al.*, *Science* **343**, 1241484 (2014).
6. T. Levi, A. M. Kilpatrick, M. Mangel, C. C. Wilmers, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 10942 (2012).

ACKNOWLEDGMENTS

The findings and conclusions are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service.

10.1126/science.abc7060

MOLECULAR BIOLOGY

CRISPR at lightning speeds

Breaking the genome with the flick of a (light) switch

By **Darpan Medhi** *and* **Maria Jasin**

Fundamental studies since the mid-1990s demonstrated the power of targeting DNA double-strand breaks (DSBs). This work uncovered a plethora of DNA repair pathways and formed the basis for the revolution in genome modification (*1*). That revolution in its most powerful manifestation came about with the discovery of RNA-directed DNA-cleaving enzymes in CRISPR systems, allowing the simple rules of base-pairing to guide targeted DNA breakage. Yet the CRISPR-associated 9 (Cas9) endonuclease that induces a DSB at target sites (determined by guide RNAs) in gene editing has been delivered as a relatively blunt instrument, with little control over its activity. On page 1265 of this issue, Liu *et al.* (*2*) developed Cas9 into a precision instrument that is both temporally and spatially controlled. It should now be possible to interrogate the cellular response to DSBs in real time, which will allow greater understanding of how cells maintain genome integrity when faced with these potentially catastrophic lesions.

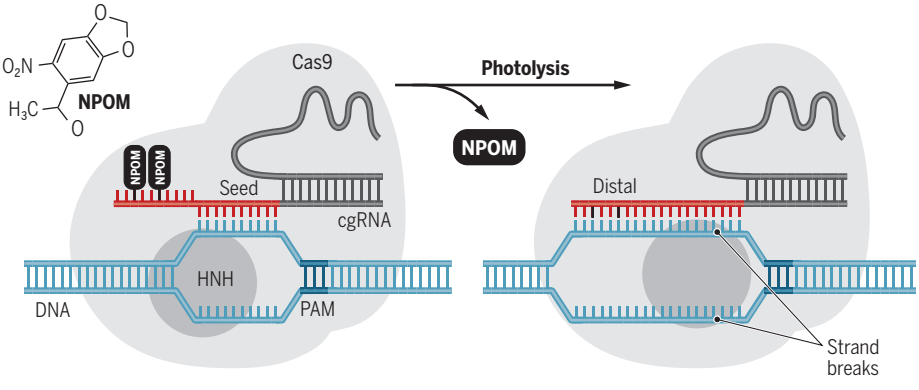
Gene-editing nucleases have historically been expressed in cells from plasmid vectors that rely on several steps (cellular uptake, transcription, and translation) before cleavage can occur. Once expressed, the nucleases may be present for days, or even permanently, if viral transduction is used. With the advent of CRISPR, direct delivery of active but relatively short-lived nucleases, in the form of Cas9–guide RNA ribonucleoproteins (RNPs) (*3*), has been feasible. However, even with RNP delivery, cleavage of genomic sites is not immediate and may not be synchronous (*3*), hampering the study of DNA breakage and repair in real time.

To overcome these limitations, Liu *et al.* developed very fast CRISPR (vfCRISPR) to enable precise control of Cas9 activity using light. Prebound to its target in genomic DNA through a modified guide RNA, Cas9 is activated to cleave DNA nearly instantaneously by shining light at wavelengths that are not phototoxic to cells (365 or 405 nm). The system uses standard CRISPR features, including protospacer adjacent motif (PAM) recognition by Cas9 and guide RNA binding close to the PAM (see the figure). However, distal guide RNA binding is blocked as a result of steric hindrance by modified deoxyribonucleotides (*4*) incorporated into the guide RNA. This "caged" guide RNA (cgRNA) effectively prevents complete R-loop (DNA-RNA hybrid) formation of the cgRNA with the target DNA (*2*), which is necessary to activate

Developmental Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA.
Email: m-jasin@ski.mskcc.org

Photoactivation of CRISPR

After protospacer adjacent motif (PAM) recognition by CRISPR-associated 9 (Cas9), the caged guide RNA (cgRNA) forms base pairs with the target DNA at the PAM-proximal "seed" sequences. However, distal cgRNA sequences are prevented from binding by steric hindrance from 6-nitropiperonyloxymethyl (NPOM) thymidine. Photolysis of NPOM (365 or 405 nm) allows distal base pairing, provoking a conformational change of the Cas9 HNH endonuclease domain. This activates DNA cleavage by both nuclease domains.



GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

the Cas9 nuclease (*5*, *6*). Because the modified residues are photolabile, unmasking occurs rapidly upon light exposure to allow full DNA base pairing.

Light has long been recognized as a desirable agent for controlling all sorts of biological processes because it is noninvasive and provides exquisite spatial and temporal control (*7*, *8*). Methods of photoactivatable Cas9 cleavage were previously reported by other groups using diverse approaches (caged, split, and shielded Cas9s; protected guide RNAs) (*9*). However, the cgRNA design demonstrates the fastest kinetics thus far.

As with many cellular processes, the DNA damage response exhibits fast kinetics and involves a myriad of factors that act in a highly choreographed fashion (*10*). With vfCRISPR, Liu *et al*. found that they could achieve cleavage of half of the DNA molecules within seconds of light exposure. Mutagenesis at the cleaved sites was strictly light-dependent and efficient. With this precise control of Cas9 cleavage, the authors examined repair kinetics at various genomic sites. Modeling of repair kinetics at one target site unexpectedly suggested repetitive cleavage and repair, possibly because this site is particularly vulnerable to one type of repair product. Moreover, with time-resolved chromatin immunoprecipitation, the recruitment and retention of components of the DNA damage signaling and repair machinery could be tracked within minutes, including the rapid responder MRE11 (*11*). Spreading of a commonly used marker of DSBs, phosphorylated histone H2AX (γH2AX) (*12*), could be resolved into two rates, one estimated at ~150 kb/min and a surprising second layer at ~460 kb/min, reaching an astounding 30 Mb from the DSB within 1 hour.

Other methods to study DNA repair are less controllable. For example, ionizing radiation has long been used to follow the fast recruitment of DNA damage response factors into cytologically visible foci (*11*, *12*). However, the damage sites are not known and cannot be readily tracked over time. Moreover, ionizing radiation induces a number of different types of damage in addition to DSBs (*13*). An alternative approach has involved the hormone induction of a restriction enzyme that only makes DSBs; however, induction is not as fast as with light, and cells are not viable because of the large number of DSBs that are induced (*14*).

One elegant application of vfCRISPR involves fluorescence imaging of genomic loci (*2*). Using a Cas9–green fluorescent protein (GFP) fusion, Liu *et al*. were able to track DSBs at both alleles of a gene. They noticed that each allele had a focus of the DNA repair protein TP53-binding protein 1 (53BP1) that underwent cycles of expansion and dissolution over time, possibly due to cycles of breakage and repair. Although the duration of the initial 53BP1 cycle differed in different cells, it was well correlated between the two alleles in a single cell. Further investigation will be necessary to determine what in the cellular milieu is responsible for the correlated behavior of 53BP1.

By cytologically marking both alleles of a gene with the Cas9-GFP fusion, the authors were able to focus light onto a single allele to achieve a high degree of spatial control over DNA cleavage through the cgRNA, leading to mutagenesis of just one gene allele. Heterozygous mutations are often difficult to achieve, yet desirable when attempting to model dominant mutations or generate mouse lines for which homozygous mutations are lethal. Presumably, the focused application of light will also prevent off-target activity.

Light-directed Cas9 activation is expected to be transformative for understanding the kinetics of the cellular response to DSBs. The exquisite spatial control of DSB formation by light could also be exploited to illuminate repair pathways in different nuclear compartments. It also provides flexibility in experimental design to introduce DSBs at multiple locations individually, sequentially, or simultaneously. Within organisms, adapted optogenetics (light-activated protein manipulation) in neurons could allow precise interrogation not only of gene function, but also of the role of DNA breaks in neuronal activity and repair pathways. Moreover, Cas9 and other CRISPR proteins have been repurposed beyond DSB induction, including for the introduction of other types of DNA lesions. How well light activation through cgRNAs translates to studying other processes remains to be seen, but there is promise for precise interrogation of other DNA repair pathways. ■

REFERENCES AND NOTES

1. M. Jasin, J. E. Haber, *DNA Repair* **44**, 6 (2016).
2. Y. Liu *et al.*, *Science* **368**, 1265 (2020).
3. S. Kim *et al.*, *Genome Res.* **24**, 1012 (2014).
4. H. Lusic, D. D. Young, M. O. Lively, A. Deiters, *Org. Lett.* **9**, 1903 (2007).
5. S. H. Sternberg, B. LaFrance, M. Kaplan, J. A. Doudna, *Nature* **527**, 110 (2015).
6. D. W. Taylor, *Nat. Struct. Mol. Biol.* **26**, 669 (2019).
7. S. Yao *et al.*, *Nat. Methods* **17**, 422 (2020).
8. K. Deisseroth, *Nat. Methods* **8**, 26 (2011).
9. S. A. Gangopadhyay *et al.*, *Biochemistry* **58**, 234 (2019).
10. M. Lisby, J. H. Barlow, R. C. Burgess, R. Rothstein, *Cell* **118**, 699 (2004).
11. O. K. Mirzoeva, J. H. Petrini, *Mol. Cell. Biol.* **21**, 281 (2001).
12. E. P. Rogakou, C. Boon, C. Redon, W. M. Bonner, *J. Cell Biol.* **146**, 905 (1999).
13. A. C. Vitor, P. Huertas, G. Legube, S. F. de Almeida, *Front. Mol. Biosci.* **7**, 24 (2020).
14. J. S. Iacovoni *et al.*, *EMBO J.* **29**, 1446 (2010).

ACKNOWLEDGMENTS

M.J. acknowledges support from NIH RM1 HG009490 and the DeWitt Wallace Basic Science Fund (D.M.).

10.1126/science.abc3997

ELECTROCHEMISTRY

Electrification of the chemical industry

Performing chemical synthesis with renewable electricity can reduce carbon emissions

By **John L. Barton**

Curbing carbon emissions while maintaining quality of life is a global challenge for manufacturing processes that will require process innovation. One approach is replacing energy from the burning of carbon-based fuels with energy supplied by "green" electrons. This goal can be achieved in some cases by simply replacing heat supplied by combustion with electrical heating (*1*). In chemical synthesis, it can also more elegantly supply reaction energy through electrochemistry. On page 1228 of this issue, Leow *et al*. (*2*) propose an electrochemical route to ethylene oxide (EO) and propylene oxide (PO) that promises cleaner, more efficient, and more selective processing. Ethylene and propylene were epoxidized electrochemically to EO and PO, respectively, at industrially relevant current densities with Faradaic (electron-specific) selectivities ~70% to the target epoxide (*2*).

Leow *et al*. coupled an electrochemical flow cell to homogeneous reactions for an overall reaction, $C_2H_2 + H_2O \rightarrow C_2H_2O + H_2$, for EO synthesis (see the figure). Two electrochemical reactions drive this reaction. Chlorine evolution occurs at the anode, $2Cl^- \rightarrow 2e^- + Cl_2$, and hydrogen evolution occurs at the cathode, $2H_2O + 2e^- \rightarrow H_2 + 2OH^-$, where e is the charge on the electron. These reactions are not particularly interesting; what is innovative is coupling these two simple reactions with three subsequent, homogeneous chemical reactions. Dissolved chlorine in the anodic solution dissociates into hydrochloric and hypochlorous acid (HCl and HOCl, respectively). The latter reacts with ethylene to make 1-chloro-2-hydroxyethane ($HOCH_2CH_2Cl$), or ethyl-

Dow Chemical Company, Midland, MI 48667, USA.
Email: jlbarton@dow.com

ene chlorohydrin. The cathodic solution is enriched in hydroxide from the H_2 evolution. The EO and Cl^- are produced when the anodic aqueous ethylene chlorohydrin solution is mixed with the basic cathodic solution. A similar route can produce PO from propylene. The authors also develop a process to produce ethylene from CO_2 and subsequently upgrade it to EO.

Technoeconomic analysis by Leow *et al.* suggests that this process could scale to produce EO at a cost comparable with current industrial practices with a lower carbon footprint when supplied with renewable energy (*2*). Such a process would be a carbon-negative path to an important, large-scale commodity chemical. Improvements are still possible, particularly in product selectivity and catalyst selection. Nonetheless, the electrochemical productivity of EO reported in this study is a factor of 10 higher than that of the electrochemical process of Simmrock and Hellemanns (*3*).

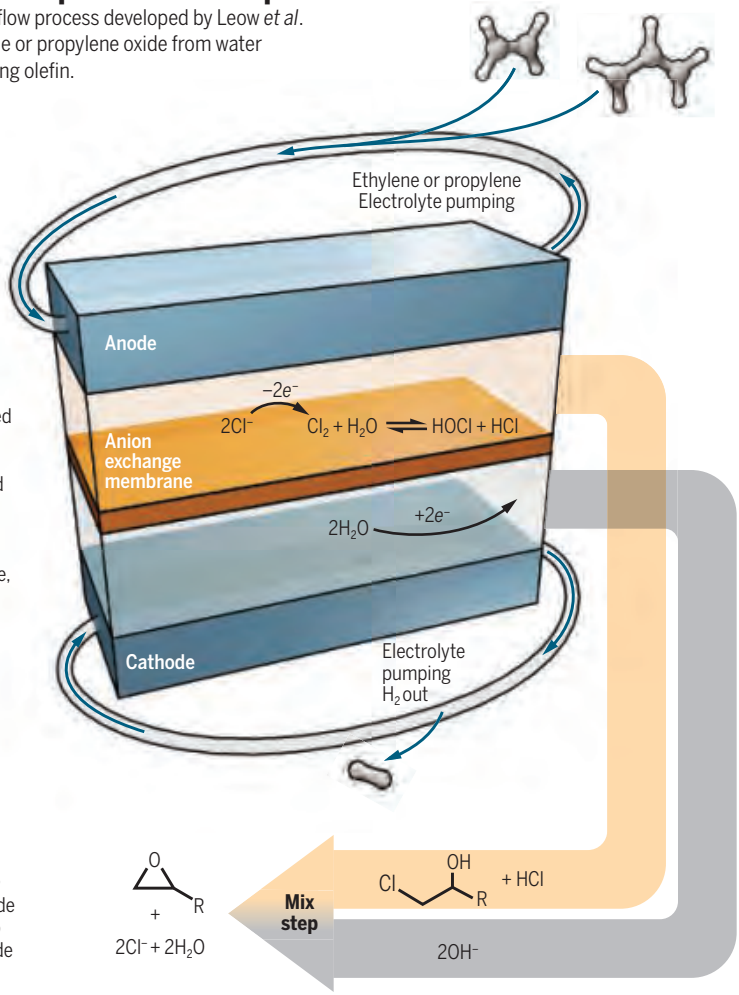
Current routes to EO and PO operate at capacities for each greater than 8 million tonnes per year (*4*, *5*). Historic production of EO was through the chlorohydrin process, which is based on chlorine produced electrochemically. This process was replaced by the direct partial oxidation process with silver catalysts, and currently, all EO is produced this way (*5*). More than one-third of global PO production still uses the chlorohydrin process, and no viable direct oxidation process with molecular oxygen is available. Most PO is now produced by using a peroxide as the oxidant (*4*). The chlorohydrin process produces immense quantities of wastewater. As new plants are built, the use of a peroxide-based process is generally more favorable financially and environmentally. The chlorohydrin process uses Cl_2 produced through the electrochemical chloro-alkali process (*6*) that operated at a lower current density than the process presented by Leow *et al.* This new process option raises the possibility that chloro-alkali assets could be repurposed to produce PO with greater efficiency and with a lower barrier to implementation.

Energy intensity, which is the total energy required to produce a product from raw materials (embodied energy), is a primary metric for comparison of process sustainability (*7*). Leow *et al.* produce EO or PO with an electrical energy demand of 0.83 MJ/mol (19 MJ/kg EO or 14 MJ/kg PO) from the corresponding olefin and water, which compares poorly with process energy estimates for direct oxidation to EO (4.0 MJ/kg EO) or a peroxidation route to PO (6.0 MJ/kg PO) (*8*). The direct oxidation route for EO will be substantially less energy intensive even if renewable electricity supplies the electrochemical process. In the case of PO, this process could be expected to have an energy input similar to that of the current chlorohydrin process, which is estimated to be more energy intensive than the hydrogen peroxide route by ~50% (*4*).

The energy intensity can be reduced through cell design (for example, changing the flow rate or electrode thickness), but the pairing of electrochemical reactions (Cl_2 and H_2 evolution) is unlikely to allow for the energy intensity to be reduced below the chloro-alkali process because it is directly proportional to the cell voltage. To make a marked improvement on the energy demand, the electrochemistry would need to be altered to reduce the cell voltage, but there could be more subtle advantages outside of the analysis here, such as a reduced water demand. Although this particular process reported by Leow *et al.* is unlikely to be the next major route to PO or EO, the development of alternative processes to commodity chemicals, especially those with clear options for incorporation of renewable energy, is necessary to realize greener processes. ■

Electrochemical production of epoxides

An electrochemical flow process developed by Leow *et al.* can produce ethylene or propylene oxide from water and the corresponding olefin.



Making the reactants
At the anode, dissolved olefin is converted to the chlorohydrin, and hydroxide is produced at the cathode along with hydrogen gas, which is vented and captured. For ethylene, R is H, and for propylene, R is CH_3.

Mixing to make products
The cathode and anode contents are mixed to produce ethylene oxide or propylene oxide.

GRAPHIC: A. KITTERMAN/*SCIENCE*

REFERENCES AND NOTES

1. H. Thunman *et al.*, *Sustainable Mater. Technol.* **22**, e00124 (2019).
2. W. R. Leow *et al.*, *Science* **368**, 1228 (2020).
3. K. H. Simmrock, G. Hellemann, U.S. Patent 4,119,507 (1978).
4. H. Baer, M. Bergamo, A. Forlin, L. H. Pottenger, J. Lindner, in *Ullmann's Encyclopedia of Industrial Chemistry* (2012); https://doi.org/10.1002/14356007.a22_239.pub3
5. S. Rebsdat, D. Mayer, in *Ullmann's Encyclopedia of Industrial Chemistry* (2001), pp. 547–568.
6. P. Schmittinger *et al.*, in *Ullmann's Encyclopedia of Industrial Chemistry* (2011); https://doi.org/10.1002/14356007.a06_399.pub3
7. J. Schwarz, B. Beloff, E. Beaver, *Chem. Eng. Prog.* **2002**, 58 (2002).
8. U.S. Department of Energy (DOE), Energy and Environmental Profile of the U.S. Chemical Industry (DOE, 2000).

10.1126/science.abb8061

GEOPHYSICS

Mapping Earth's deepest secrets

Sequencing seismograms pinpoint new structures near Earth's core-mantle boundary

By **Meghan S. Miller**

Deep within Earth's interior, at ~2900 km beneath the surface, lies the boundary between the solid silicate rock mantle and the liquid iron-nickel alloy core (the core-mantle boundary). Geophysicists have studied the complex thermal and chemical dynamics that take place in this boundary layer. In the early 20th century, Gutenberg investigated the structure of the lowermost region, or base, of the mantle by recording with only a few seismograms from a small number of large-magnitude earthquakes that occurred thousands of kilometers away (*1*). The structure of the rocks just above the core-mantle boundary—designated as D′′ by Jeffreys in 1939 (*2*)—forms a distinct layer with surprising complexity. Now, on page 1223 of this issue, Kim *et al.* (*3*) describe new structural heterogeneities in the lowermost mantle with the use of a learning algorithm that does not require any a priori knowledge of Earth.

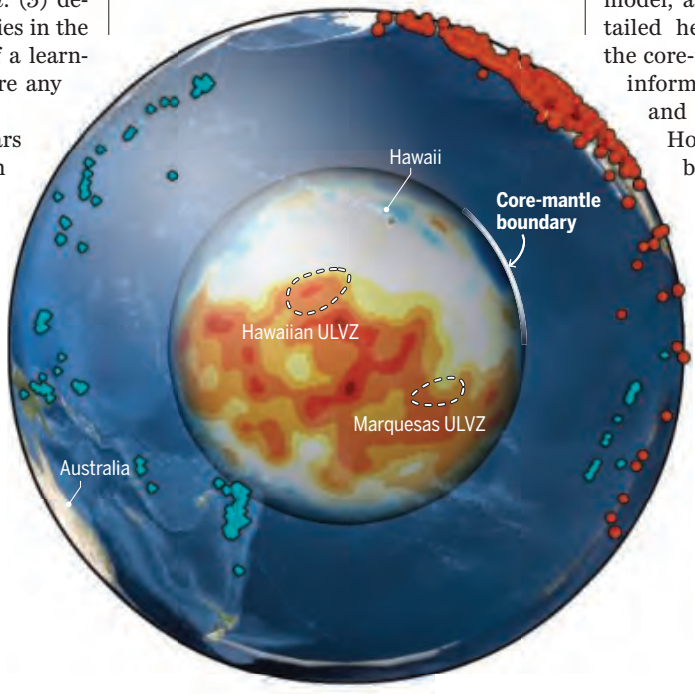
Research over the past ~100 years has yielded major improvements in scientists' understanding of the lowermost mantle. However, the velocity discontinuities (which represent the boundaries between layers within the deep Earth) detected by the pioneers in seismology with just a small number of measurements remain as fundamental constraints of Earth's structure.

Thanks to community experiments such as EarthScope (*4*), the relatively disparate and important observations of the lowermost mantle have increased in number and location with the exponential growth in seismic data collected over the past couple of decades. Seismologists have used seismograms of earthquakes recorded by arrays of distant seismometers to image the deep-mantle structures (*5*). The development of methods, such as those used by Kim *et al.*, to process and analyze increasingly large datasets are crucial to improving geophysicists' knowledge of Earth's structure, which is central to understanding the evolution of Earth.

Global tomographic models at lower mantle depths, such as S40RTS (see the figure) (*6*), were generated from low-resolution (hundred- to thousand-kilometer-scale) images of Earth's structure; such models were hampered by the use of only long-wavelength seismic waves and the type of imaging method. Despite the limitations, these global models illustrate the remarkable complexity of the lowermost mantle, which suggests that this region is one of Earth's most important thermal and structural zones. This part of Earth's interior includes large low-shear-velocity provinces (LLSVPs) and ultralow-velocity zones (ULVZs), which have been linked to hotspot volcanoes and large igneous provinces at Earth's surface (*7*, *8*). The extent and edges of structures that cluster primarily at the edges of LLSVPs [such as the African LLSVP (*9*), the Pacific LLSVP (*10*), and in particular the small-scale ULVZs (*5*, *11*)] traditionally have been mapped by analyzing individual seismic waveforms that sample the deepest mantle.

Only certain seismic phases are sensitive to the LLSVP and ULVZ structures; these include seismic waves that are diffracted off the core-mantle boundary, such as S_{diff} (*6*) and P_{diff} (*12*, *13*), or are reflected at the core-mantle boundary, such as *ScS* or *ScP* (*13*). The waveform distortion of these seismic phases, paired with arrival-time anomalies relative to a preexisting model, allows for inference about the detailed heterogeneities of the rocks near the core-mantle boundary and provide key information regarding the composition and dynamics of the lower mantle. However, after selecting data on the basis of signal-to-noise ratios and sparse earthquake and seismometer (source-receiver) geometries, the seismic waveforms often are analyzed manually and then compared with either one-dimensional (1D) models or 3D models (from tomographic imaging) of Earth's structure. Although there are millions of earthquake records available from seismic networks across the globe, it remains a challenge to analyze and quantitively assess individual seismic waveform data and extract the desired information. The ability to detect and analyze subtle changes in waveforms is important for providing constraints on the physical parameters of deep Earth.

To this end, Kim *et al.* used a manifold algorithm called "the Sequencer" to investigate a relatively large S_{diff} waveform dataset; the Sequencer enabled a data-driven analysis of the deep mantle without any expectations or prior knowledge



Shaking things up at the center of Earth

Seismic-wave scattering can pinpoint structures beneath Earth's surface, as shown in this tomographic model of Earth and its lower-mantle depths near the core-mantle boundary (center circle). The red-orange region indicates the Pacific large low-shear-velocity province (LLSVP). Green symbols indicate deep-focus (>150 km depth) earthquakes that occurred between 2000 and 2018 and were recorded at distant seismometers (red symbols). The seismograms were analyzed by Kim *et al*. ULVZ, ultralow velocity zone.

GRAPHIC: ADAPTED BY V. ALTOUNIAN/*SCIENCE* FROM NATIONAL COMPUTATIONAL INFRASTRUCTURE AUSTRALIA VIZLAB

Research School of Earth Sciences, Australian National University, Canberra, Australia. Email: meghan.miller@anu.edu.au

about its structure. This unsupervised, graph-based algorithm orders the waveforms to minimize dissimilarities and can reveal trends without an Earth model. The analysis by Kim *et al.* detected subtle changes in seismic waveforms from earthquakes that occurred in Asia and Oceania and were recorded in the Americas and mapped their origins across a large geographic region beneath the Pacific Ocean.

The new study also identified more broadly distributed ULVZs at the base of the mantle north of Hawaii and a previously undetected anomaly in the deepest mantle beneath the south-central Pacific. This type of analysis could be applied to various seismic phases such as *ScS*, *ScP*, and P_{diff} and a range of others that are of higher frequency, which would provide a new, higher-resolution, and more comprehensive mapping of the structural heterogeneity of deep Earth. Knowledge of these physical properties and of inferred chemical and thermal structures is essential to determining whether partial melt of the rocks exists at the core-mantle boundary, whether distinct materials accumulate or stabilize in particular regions, whether some volcanoes have origins in deep Earth, and, last, what the compositional variations are in the lowermost mantle. ■

"... a manifold algorithm... enabled a data-driven analysis of the deep mantle without any expectations or prior knowledge about its structure."

REFERENCES AND NOTES

1. B. Gutenberg, *Nachr. Ges. Wiss. Goettingen Math. Phys. Kl.* **1914**, 125 (1914).
2. H. Jeffreys, *Mon. Not. R. Astron. Soc. Geophys.* **4** (suppl.), 498 (1939).
3. D. Kim *et al.*, *Science* **368**, 1223 (2020).
4. EarthScope Working Group, *Eos* **81**, 122 (2000).
5. S. Yu, E. J. Garnero, *Geochem. Geophys. Geosyst.* **19**, 396 (2018).
6. J. Ritsema *et al.*, *Geophys. J. Int.* **184**, 1223 (2011).
7. K. Burke *et al.*, *Earth Planet. Sci. Lett.* **265**, 49 (2008).
8. Q. Williams *et al.*, *Science* **281**, 546 (1998).
9. S. Ni *et al.*, *Science* **296**, 1850 (2002).
10. Y. He, L. Wen, *J. Geophys. Res. Solid Earth* **117**, B09308 (2012).
11. E. J. Garnero, D. V. Helmberger, *J. Geophys. Res. Solid Earth* **103**, 12495 (1998).
12. D. A. Frost, S. Rost, *Earth Planet. Sci. Lett.* **403**, 380 (2014).
13. S. Rost *et al.*, *Nature* **435**, 666 (2005).

ACKNOWLEDGMENTS

I thank the National Computational Infrastructure Australia Vizlab for assistance with the figure.

10.1126/science.abc3134

ECOLOGY

The global odyssey of plastic pollution

Thinking big about small particles reveals new features of the microplastic cycle

By **Chelsea M. Rochman**[1] *and* **Timothy Hoellein**[2]

Scientists who once studied microplastics (plastic debris <5 mm in size) as ocean pollutants have now detected them in soils, biota, and Earth's atmosphere. To decipher the global fate of microplastics, scientists have begun to ask questions about the "microplastic cycle," which is akin to global biogeochemical cycles (nitrogen, carbon, and water). For example, what are the sources of microplastics, and how do they transform as they move from one pool (e.g., a beach, inside an organism, or a river bed) to another? And what processes ("fluxes") transfer microplastics between pools? On page 1257 of this issue, Brahney *et al.* (*1*) report high-resolution spatial and temporal data that provide evidence of both global and regional microplastic transport, thus increasing our understanding of the microplastic cycle.

Nearly a decade ago, scientists began studying marine microplastics in surface currents of the ocean as the key mechanism for global transport. As datasets grew, their understanding of long-range transport within and between oceans expanded to include mechanisms such as deep-sea circulation (*2*), biological transport (*3*, *4*), and drifting sea ice (*5*). In parallel, emerging work uncovered pools of microplastics in other Earth compartments, including freshwater and terrestrial systems (*6*), and the atmosphere (*7*). To fully understand the microplastic cycle, researchers must piece together the fluxes that connect the transport and transformation of microplastics as they move between planetary compartments.

Atmospheric transport of microplastics in airborne dust, which settles to the ground (deposition) during dry and wet periods and in both urban and remote locations, was initially overlooked. Scientists have understood the global transport of dust for decades, but until recently, dust was not known to carry substantial amounts of microplastics. Seminal work on transport of microplastics in the atmosphere demonstrated their presence in wet (e.g., rain, snow) and dry deposition in Paris, France (*8*), providing proof that microplastics are a component of dust and that atmospheric deposition is a mechanism of transport. Long-range atmospheric transport, away from urban centers, was first demonstrated in 2019, when microplastics were unearthed from a remote mountain catchment (*7*) and in Arctic snow (*9*). These studies prove that microplastics are transported atmospherically to both regional and faraway places.

Brahney *et al.* studied both global and regional transport of microplastics by comparing the size and shape of particles deposited in dry and wet weather (see the figure). The new work elucidates patterns of deposition and processes of atmospheric transport, and predicts that atmospheric transport is an important source of microplastics in remote locations. For example, the authors estimate that more than 122 tons of microplastics are deposited annually to U.S. protected lands of the western United States.

By incorporating human population metrics, local weather patterns, and climate models, Brahney *et al.* found that larger microplastics were deposited during wet events and likely originated from nearby urban centers during regional storms. In contrast, smaller microplastics deposited during dry weather were more likely to have been transported long distances and made up the majority of the microplastic mass.

A key insight from the new work is that fundamental tools for studying global dust transport can be applied to microplastics. Like dust, most particles measured were within the size range typical of global transport (<25 μm). However, microplastics are less dense than soil and therefore might travel longer distances than natural dust particles. Future research should test hypotheses about the distances that microplastics can travel atmospherically and the processes that entrain microplastics in the air, such as sea spray and dust storms. The new study also invites questions about latitudinal gradients. Global atmospheric circulation is affected, in part, by air rising at 0° and 60° latitude and

[1]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada. [2]Department of Biology, Loyola University Chicago, Chicago, IL, USA. Email: chelsea.rochman@utoronto.ca

sinking at 30° and 90°. Do researchers expect movement of atmospheric microplastics to follow similar patterns? Moreover, scientists know that dust particles are a vector for atmospheric distribution of chemical contaminants and microbes. How do microplastics interfere or aid in such spread?

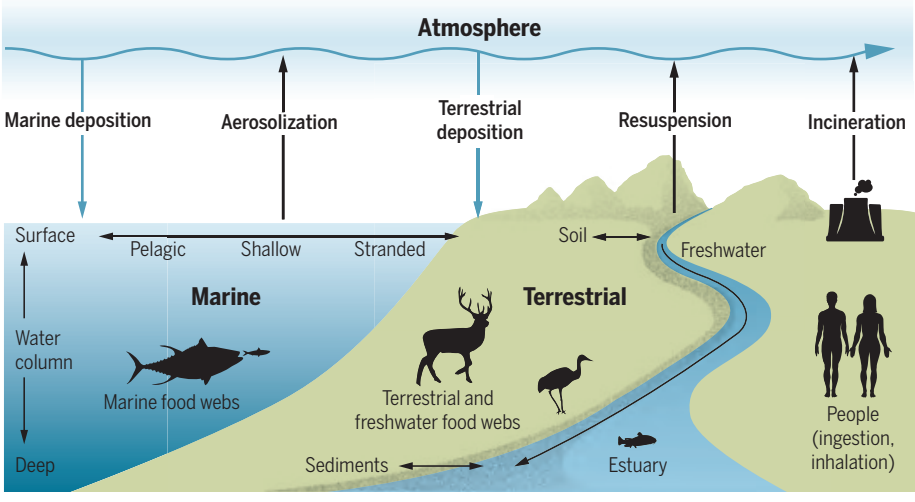
Researchers also must delve deeper into other transport pathways of the microplastic cycle. For example, scientists know little about how microplastics move through terrestrial ecosystems, but evidence suggests percolation from the surface to deeper layers, resuspension into the atmosphere, and transport into ground and surface waters (*9*). In streams, microplastics can be continuously deposited and resuspended, buried in sediments, or exported to downstream ecosystems including lakes, estuaries, and the ocean (*10*). Also, very little is known about how microplastics move through food webs, but some evidence suggests trophic transfer (*11*). Moreover, during their life cycle, microplastics are subject to abiotic and biotic transformations. For example, microbes can incorporate carbon atoms derived from plastic when building their cell membranes (*12*). Microbial breakdown can generate airborne fragments and greenhouse gases, as can degradation (oxidation) by ultraviolet light from the Sun (*13*).

Researchers now recognize that plastic cycles through more than just the ocean. To be sure, the ocean is a sink for microplastics. But as ecologists learn more about these particles, they realize that the ocean is not always the final fate of microplastics; rather, the ocean is one part of the global plastic cycle. In his essay "Odyssey," Aldo Leopold poetically describes how biological, chemical, and geological factors shape the journey of atoms "X" and "Y" (representing carbon and nitrogen) among the land, atmosphere, and sea, between biotic and abiotic compartments, and through deep time. Microplastics might display a different arrangement of atoms, but nonetheless, plastic particles in the environment are subject to the same biogeochemical forces as all other matter.

As ecologists think about all that encompasses the sources, fates, and transformations of microplastics, Leopold's prose challenges them to envision and measure processes among all ecosystem compartments, including those that strain one's imagination, such as plastic in rain. Future microplastic researchers should follow the example of Brahney *et al.* and think big about small particles, so as to contribute to our understanding of the global microplastic cycle and how it has become an aspect of global elemental cycles. ■

Microplastic pollution is pervasive

Emerging research pinpoints atmospheric deposition as a mode of microplastic transfer to the western United States. Mapping microplastic pools (water, land, organisms) and fluxes (arrows) will guide delineation of the global microplastic cycle.



GRAPHIC: N. CARY/*SCIENCE*

REFERENCES AND NOTES

1. J. Brahney *et al.*, *Science* **368**, 1257 (2020).
2. I. A. Kane *et al.*, *Science* **368**, eaba5899 (2020).
3. K. Katija *et al.*, *Sci. Adv.* **3**, e1700715 (2017).
4. A. Porter *et al.*, *Environ. Sci. Technol.* **52**, 7111 (2018).
5. I. Peeken *et al.*, *Nat. Commun.* **9**, 1505 (2018).
6. A. A. Horton *et al.*, *Sci. Total Environ.* **586**, 127 (2017).
7. S. Allen *et al.*, *Nat. Geosci.* **12**, 339 (2019).
8. R. Dris *et al.*, *Environ. Chem. Lett.* **12**, 592 (2015).
9. M. Bergmann *et al.*, *Sci. Adv.* **5**, eaax1157 (2019).
10. T. J. Hoellein *et al.*, *Sci. Rep.* **9**, 3740 (2019).
11. J. F. Provencher *et al.*, *Environ. Rev.* **27**, 304 (2019).
12. S. J. Taipale *et al.*, *Sci. Rep.* **9**, 19894 (2019).
13. B. Gewert *et al.*, *Environ. Sci. Process. Impacts* **17**, 1513 (2015).

ACKNOWLEDGMENTS

We thank M. J. Hoffman and D. L. Mahler for comments.

10.1126/science.abc4428

GENETICS

A gene for color differences between sexes

Sex differences in plumage color in hybrid canaries are controlled by a single enzyme

By **Nancy Chen**

Sexual dimorphism—phenotypic differences between sexes of the same species—is a widespread yet puzzling phenomenon in nature. How such traits evolve has fascinated evolutionary biologists since Darwin, whose ponderings about elaborate ornaments in males prompted him to develop the theory of sexual selection (*1*), and numerous studies have sought to explain the evolution of these traits (*2*). Less is known about the genetic and molecular mechanisms that allow species to generate divergent morphologies from nearly identical genomes (*3*). On page 1270 of this issue, Gazda *et al.* (*4*) show that sex-specific plumage coloration in hybrid canaries is controlled by a single genomic region containing the gene encoding β-carotene oxygenase 2 (*BCO2*). Differences in coloration between the sexes are due to the up-regulation of *BCO2* expression and consequent degradation of pigments in females, demonstrating that color differences between males and females can evolve through a simple molecular mechanism.

One of the most prominent forms of sexual dimorphism is different coloration between males and females, or sexual dichromatism. Indeed, in some species, sex-specific coloration is so divergent that males and females were once considered different species (*1*). Sexual dichromatism is typically thought to arise from sexual selection for increased ornamentation in males (*1*) and, accordingly, is often used as a proxy for the strength of sexual selection in comparative studies (*5*). Alternatively, sexual dichromatism could also arise from natural selection for cryptic coloration in females (*6*). Recent studies indicate that sexual dichromatism is driven by sexual and natural selection operating in both

Department of Biology, University of Rochester, Rochester, NY, USA. Email: nancy.chen@rochester.edu

sexes (*7*, *8*). A better understanding of the genetic basis of sexual dichromatism is crucial for fully elucidating the underlying evolutionary processes.

In birds, sexual dichromatism has evolved rapidly, with multiple gains and losses of dichromatism among closely related species (*7*), and is primarily driven by differences in carotenoid-based coloration (*9*). Carotenoids are the pigments responsible for most red, orange, and yellow colors observed in nature (*10*). These pigments are primarily synthesized by plants, bacteria, and fungi; most animals have to acquire carotenoids through their diet (*10*). Carotenoid-based coloration is widely considered to be a signal of individual quality or condition, although the exact physiological mechanisms of carotenoid signaling remain contentious (*11*). How carotenoids are processed and deposited in peripheral tissues and how sex-specific differences in carotenoid-based coloration arise are largely unclear.

Gazda *et al*. studied the genetic basis of sexual dichromatism in domesticated canaries. They focused on *mosaic* canaries, which are generated by crossing the sexually dimorphic red siskin (*Spinus cucullatus*) with the sexually monomorphic common canary (*Serinus canaria*), followed by repeatedly crossing hybrid progeny with *S. canaria* (backcrossing). *Mosaic* canaries can be yellow or red, and males exhibit more carotenoid pigmentation than females (see the figure). On the basis of the breeding design, *mosaic* canaries should be genetically similar to the common canary throughout their genome except at the regions responsible for sexual dichromatism (the *mosaic* phenotype), which should be derived from the red siskin genome. By analyzing canary genome sequences, Gazda *et al*. identified a single 32-kb region of the genome associated with sexual dichromatism.

To further narrow down their search, Gadza *et al*. studied gene expression patterns. All sexually dimorphic traits are generated by sex differences in gene expression during development, and most are controlled by the sex hormones estrogen and testosterone (*3*). The authors identified a single gene in the genomic region of interest that was differentially expressed in the feather follicles of males and females: *BCO2*. The encoded enzyme degrades carotenoid molecules and is associated with variation in carotenoid concentrations in sheep, cows, and birds (*10*). Concordant with previous studies, Gazda *et al*. show that *BCO2* is selectively expressed in developing white feather follicles, concluding that the reduced carotenoid pigmentation in *mosaic* females is caused by sex-specific upregulation of *BCO2*. These results provide evidence that sex differences in carotenoid pigmentation arise not from differential deposition of pigments but from differential local degradation of pigments caused by sex-biased expression of *BCO2*. This observation is intriguing because recent avian phylogenetic studies show that variation in sexual dichromatism is often driven by the loss of bright colors in females, contrary to the common assumption that dichromatism evolves through the gain of bright, elaborate plumage in males (*8*).

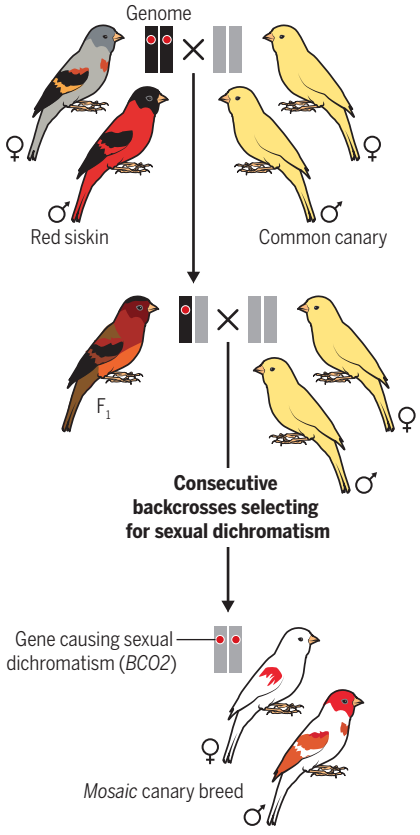
Does *BCO2* also explain sexual dichromatism in other species? To test the generality of their mechanism, Gazda *et al*. examined gene expression in developing feathers of three species with varying degrees of sexual dichromatism in carotenoid-based pigmentation: the common canary (*S. canaria*), the European serin (*Serinus serinus*), and the house finch (*Haemorhous mexicanus*). Patterns of *BCO2* expression in the European serin are consistent with patterns observed in the *mosaic* canaries: Females and feather patches with less carotenoid pigmentation have higher expression of *BCO2*. By contrast, *BCO2* expression was not associated with carotenoid pigmentation in house finches, suggesting that multiple mechanisms for sexual dichromatism exist even within finches.

The findings of Gazda *et al*. provide an important contribution to the understanding of the genetic basis of sexual dichromatism and reveal several exciting avenues for future research. The authors provide evidence that suggests that the regulation of *BCO2* expression may be complex, with multiple modifiers of expression located both near the gene (cis-regulatory elements) and in more distant regions of the genome (trans-regulatory elements). Further, sexual dichromatism is known to be linked to estrogen concentrations in *mosaic* canaries, because reproductively senescent and ovariectomized females develop the same plumage as males (*4*). Exactly how estrogen affects *BCO2* expression remains an open question. The regulatory pathway of *BCO2*, as well as those of other genes involved in depositing or degrading carotenoids, is a prime candidate in understanding the mechanisms for sexual dimorphism in other taxa.

The simple genetic and molecular basis of carotenoid-based sexual dichromatism illustrated in the study of Gazda *et al*. may help explain why sexual dichromatism evolves so rapidly. Further studies of this locus may help disentangle the relative importance of natural and sexual selection, and the role of sex chromosomes, in the evolution of sexual dichromatism. Researchers are now poised to deepen the understanding of the evolutionary origin and maintenance of sexually divergent traits. ■

Canary coloration

Sex-specific colors (sexual dichromatism) in hybrid canaries are controlled by a single gene encoding β-carotene oxygenase 2 (*BCO2*) that is inherited from the red siskin. They arise by crossing and backcrossing red siskins and common canaries and selecting for sexual dichromatism.



REFERENCES AND NOTES

1. C. Darwin, *The Descent of Man, and Selection in Relation to Sex* (John Murray, 1871).
2. M. Andersson, *Sexual Selection* (Princeton Univ. Press, 1994).
3. T. M. Williams, S. B. Carroll, *Nat. Rev. Genet.* **10**, 797 (2009).
4. M. A. Gazda *et al*., *Science* **368**, 1270 (2020).
5. K. Kraaijeveld *et al*., *Biol. Rev. Camb. Philos. Soc.* **86**, 367 (2011).
6. A. R. Wallace, *Darwinism: An Exposition of the Theory of Natural Selection with Some of Its Applications* (Macmillan, 1889).
7. A. V. Badyaev, G. E. Hill, *Annu. Rev. Ecol. Evol. Syst.* **34**, 27 (2003).
8. J. J. Price, *J. Ornithol.* **160**, 1213 (2019).
9. A. V. Badyaev, G. Hill, *Biol. J. Linn. Soc. Lond.* **69**, 153 (2000).
10. D. P. L. Toews *et al*., *Trends Genet.* **33**, 171 (2017).
11. P. A. Svensson *et al*., *Behaviour* **148**, 131 (2011).

ACKNOWLEDGMENTS

N.C. is supported by NIH grant R35GM133412.

10.1126/science.abc2242

GRAPHIC: MELISSA THOMAS BAUM/*SCIENCE*

PHOTO: PHOTO12/UNIVERSAL IMAGES GROUP VIA GETTY IMAGES

Oral poliovirus vaccine (OPV), comprising live attenuated viruses, can reduce the incidence of other infections. Albert Sabin, who developed an OPV, administers the vaccine, 1966.

VIEWPOINT: COVID-19

Can existing live vaccines prevent COVID-19?

Live vaccines can prevent unrelated infections and may temporarily protect against COVID-19

By **Konstantin Chumakov[1,2], Christine S. Benn[3], Peter Aaby[4], Shyamasundaran Kottilil[5], Robert Gallo[2,5]**

Prophylactic vaccination is the most effective intervention to protect against infectious diseases. The commonly accepted paradigm is that immunization with both attenuated virus (live but with substantially reduced virulence) and inactivated (killed virus particles) vaccines induces adaptive and generally long-term and specific immunity in the form of neutralizing antibodies and/or activating pathogen-specific cellular immune responses. However, an increasing body of evidence suggests that live attenuated vaccines can also induce broader protection against unrelated pathogens likely by inducing interferon and other innate immunity mechanisms that are yet to be identified. The stimulation of innate immunity by live attenuated vaccines in general, and oral poliovirus vaccine (OPV) in particular, could provide temporary protection against coronavirus disease 2019 (COVID-19).

OPV was developed by Albert Sabin in the 1950s and consists of live attenuated polioviruses of the three serotypes. Early clinical studies showed that besides protecting against poliomyelitis, OPV reduced the number of other viruses that could be isolated from immunized children, compared with placebo recipients. Additional evidence of nonspecific effects of OPV came from the 1959 poliomyelitis outbreak in Singapore caused by type 1 poliovirus that was successfully stopped by the use of monovalent OPV that contained only type 2 poliovirus (*1*). Monovalent OPVs do not induce cross-neutralizing antibodies that target other virus serotypes, so the most plausible explanation was viral interference, which presumably is mediated by innate immunity.

Large-scale clinical studies of OPV for nonspecific prevention of diseases were carried out in the 1960s and 1970s. These involved more than 60,000 individuals and showed that OPV was effective against influenza virus infection, reducing morbidity 3.8-fold on average (*2, 3*). OPV vaccination also had a therapeutic effect on genital herpes simplex virus infections, accelerating healing. OPV not only demonstrated positive effects against viral infections but also had oncolytic properties, both by directly destroying tumor cells and by activating cellular immunity toward tumors (*2*). These observations were among the first examples of viral oncotherapy, which is being actively pursued.

To extend the therapeutic effect of OPV against viral infections and cancer, three monovalent OPVs were used sequentially. In addition, several nonpathogenic enteroviruses isolated from healthy individuals [called live enterovirus vaccines (LEVs)] were tested for safety (*2*). They were also used in more than 90,000 individuals for nonspecific protection and found to reduce incidence of seasonal influenza and acute respiratory diseases (*2*, *3*). Subsequently, in Bulgaria, mass immunization with OPV helped to control a 1975 outbreak of unrelated acute poliomyelitis-like disease caused by Enterovirus 71 (*4*).

[1]Office of Vaccines Research and Review, Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, MD, USA. [2]Global Virus Network, Baltimore, MD, USA. [3]OPEN and Danish Institute for Advanced Study, University of Southern Denmark, Odense, Denmark. [4]Bandim Health Project, Bissau, Guinea-Bissau. [5]Institute of Human Virology, University of Maryland School of Medicine, Baltimore, MD, USA. Email: konstantin.chumakov@fda.hhs.gov

More recent studies confirmed these broad protective effects of OPV. Data from a randomized controlled trial (RCT) of OPV in Guinea-Bissau, West Africa, showed that OPV given at birth reduced infant mortality by ~32% (*5*). In addition, an analysis of the effect of annual and biannual national OPV immunization campaigns showed that they reduced all-cause mortality by 19%, with each subsequent campaign adding a further 13% reduction (*6*). This means that repeated immunization has an additive effect despite antibodies induced by the first vaccination. Depending on initial age, it was necessary to give OPV to between 68 and 230 children to prevent one death within the first 3 years of life (*6*). These observations were made in the complete absence of poliovirus circulation, emphasizing the nonspecific nature of the OPV-induced protection. In RCTs comparing OPV against inactivated polio vaccine (IPV), it was found that OPV reduced the burden of bacterial diarrheal disease in infants in Bangladesh (*7*). In Finland, immunization with OPV was associated with less doctor-diagnosed acute otitis media (middle ear infection, which can be caused by both viruses and bacteria) than in the IPV-immunized group (*8*). Furthermore, a retrospective study from Denmark found that the use of OPV was associated with reduced hospital admissions for respiratory infections in children.

These nonspecific beneficial effects may not be limited to OPV and LEV. Other live attenuated viral vaccines such as those against measles (*5*) and smallpox (*9*) have also been associated with pronounced nonspecific protective effects against infectious diseases. In Africa, when measles vaccine was introduced in the community, the overall mortality in children declined by more than 50%, a reduction that was far larger than anticipated on the basis of the protection against measles deaths alone (*10*). A large-scale RCT confirmed that the measles vaccine was associated with a 30% reduction in overall mortality in children; only 4% was explained by prevention of measles infection (*5*).

Attenuated bacterial vaccines such as Bacillus Calmette–Guérin (BCG) against tuberculosis, as well as experimental live attenuated vaccine against pertussis (whooping cough), were also shown to protect against heterologous infections (*5*, *11*). In addition, live pertussis vaccine also prevented noninfectious inflammatory diseases (*11*). RCTs showed that BCG vaccine at birth was associated with more than a one-third reduction of neonatal mortality, because BCG vaccine protected against deaths from septicemia and pneumonia (*5*). In 2014, an expert panel at the World Health Organization reviewed the evidence for nonspecific effects of live vaccines and concluded that they reduced childhood mortality by more than would be expected through their effects on the diseases they prevent (*12*). It is important to note that non-live (inactivated) vaccines do not seem to have the same effects, suggesting that replicating attenuated pathogens induce a broader immune response.

Some of the nonspecific effects of vaccines may be mediated by interferons. However, the mechanism may be more complex and diverse. Numerous studies have shown that BCG activates the innate immune system, resulting in enhanced responsiveness to subsequent triggers, so-called "trained innate immunity" (*13*). BCG given 4 weeks prior to a yellow fever vaccine significantly reduced virus load, confirming that it could modify the course of a viral challenge in vivo. The effects were mediated through epigenetic modifications in innate immune cells, leading to higher innate cytokine production (*13*). BCG can also induce emergency granulopoiesis within hours of administration, leading to a marked increase in the number of circulating neutrophils, providing protection from sepsis (*14*). The duration of the nonspecific protection induced by live vaccines is unknown but has been observed to last for many months to years after vaccination. For example, BCG given at school entry (5- to 6-year-olds) in Denmark was associated with a 42% reduction in the risk of dying from natural causes until the age of 45 years (*9*).

Recent reports indicate that COVID-19 may result in suppressed innate immune responses (*15*). Therefore, stimulation by live attenuated vaccines could increase resistance to infection by the causal virus, severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2). Clinical studies of this hypothesis could begin immediately. BCG trials have already been initiated by immunizing frontline health care workers (NCT04327206, NCT04328441). The endpoint of these trials is the difference in COVID-19 incidence, duration, and severity between immunized and unimmunized populations.

We propose the use of OPV to ameliorate or prevent COVID-19. Both poliovirus and coronavirus are positive-strand RNA viruses; therefore, it is likely that they may induce and be affected by common innate immunity mechanisms. There are multiple important advantages to using OPV: a strong safety record, the existence of more than one serotype that could be used sequentially to prolong protection (*2, 3*), low cost, ease of administration, and availability. Over 1 billion doses of OPV are produced and used annually in more than 140 countries. Although the supply of BCG is limited, a small fraction of OPV intended for the suspended polio eradication campaign would be sufficient for the clinical trials, and provided a positive outcome, production could likely be scaled up quickly.

Another advantage of OPV over BCG is safety. Up to 1% of BCG recipients require medical attention, owing to adverse reactions. The risk of complications due to OPV is extremely low. Vaccine-associated paralytic polio (VAPP) develops in 1 per 3 million vaccine doses given to unimmunized individuals and mostly occurs in immunocompromised children. Sequential use of IPV followed by OPV demonstrated that prior immunization eliminates the risk of VAPP. In populations with inadequate immunity, OPV was also shown to generate circulating vaccine-derived polioviruses (cVDPVs). However, in countries with sufficient vaccine coverage, the risk is minimal: Over 35 years of OPV use in the United States has resulted in no documented case of cVDPV. Therefore, if used properly, OPV is likely a safer choice than BCG.

If the results of RCTs with OPV are positive, OPV could be used to protect the most vulnerable populations. However, OPV would be most effective if the entire population of a country or region is immunized synchronously. OPV produces herd effects, and in addition to protecting vulnerable individuals, could also prevent the spread of SARS-CoV-2 by increasing the proportion of unsusceptible individuals. The strategy of inducing nonspecific protection may even have an advantage over a SARS-CoV-2–specific vaccine if SARS-CoV-2 undergoes mutation that leads to antigenic drift (and loss of vaccine efficacy), similar to seasonal influenza viruses. If proven to be effective against COVID-19, emergency immunization with live attenuated vaccines could be used for protection against other unrelated emerging pathogens. ■

REFERENCES AND NOTES

1. J. H. Hale *et al.*, *BMJ* **1**, 1541 (1959).
2. M. K. Voroshilova, *Prog. Med. Virol.* **36**, 191 (1989).
3. M. P. Chumakov *et al.*, *Zh. Mikrobiol. Epidemiol. Immunobiol.* 37 (1992).
4. L. M. Shindarov *et al.*, *J. Hyg. Epidemiol. Microbiol. Immunol.* **23**, 284 (1979).
5. P. Aaby, C. S. Benn, *Clin. Microbiol. Infect.* **25**, 1459 (2019).
6. A. Andersen *et al.*, *Front. Public Health* **6**, 13 (2018).
7. A. Upfill-Brown *et al.*, *Clin. Infect. Dis.* **65**, 414 (2017).
8. E. Seppälä *et al.*, *Vaccine* **29**, 8615 (2011).
9. A. Rieckmann *et al.*, *Int. J. Epidemiol.* **46**, 695 (2017).
10. P. Aaby *et al.*, *BMJ* **311**, 481 (1995).
11. S. Cauchi, C. Locht, *Front. Immunol.* **9**, 2872 (2018).
12. J. P. Higgins *et al.*, *BMJ* **355**, i5170 (2016).
13. R. J. W. Arts *et al.*, *Cell Host Microbe* **23**, 89 (2018).
14. B. Brook *et al.*, *Sci. Transl. Med.* **12**, eaax4517 (2020).
15. M. Zheng *et al.*, *Cell. Mol. Immunol.* **17**, 533 (2020).

10.1126/science.abc4262

RETROSPECTIVE

Lord Robert May (1936–2020)

Pioneering ecologist and a founder of chaos theory

By **H. Charles J. Godfray**[1] *and* **Angela R. McLean**[2]

Lord Robert May, groundbreaking ecologist, died in Oxford, England, on 28 April. He was 84. In addition to ecology, Bob made substantial contributions to a wide variety of fields, including epidemiology and economics. A towering figure in British and international scientific life, he served as both chief scientific adviser for the United Kingdom and president of the Royal Society. Slight of build with a formidable intellect, Bob changed the way we investigate the dynamics of the living world.

Born in Sydney, Australia, on 8 January 1936, Bob attended Sydney Boys High School, where he excelled both academically and athletically. In 1956, he received his bachelor's degree in physics from the University of Sydney. He stayed on to earn his Ph.D. in superconductivity in 1959 and then conducted further research in theoretical physics as a faculty member.

In the late 1960s, Bob learned of a movement in ecology, associated with George Evelyn Hutchinson and Robert MacArthur, that aimed to put the subject on a more rigorous and quantitative footing. At the time, some of the simplest models in ecology, such as those describing predator-prey dynamics, were often unstable, leading to predictions of the extinction of some or all of the modeled species. It was tacitly assumed that when these simple systems were embedded in more complex ecosystems, stability would inevitably emerge, but there was no way to assess this assumption. In his 1973 monograph, *Stability and Complexity in Model Ecosystems*, Bob showed that stability was not inevitable. He introduced new mathematical techniques to ecology and initiated a research program that to this day continues to investigate what aspects of natural communities allow their persistence. So impressive were Bob's immediate contributions to ecology that, upon MacArthur's tragic and untimely death in 1972, Bob was offered his position in ecology at Princeton University.

Models of a single species competing for access to food often predict a transition from stable to cyclic population dynamics as birth rate increases. Bob studied these models and found that as birth rate continues to increase, more complicated patterns of cycles occur, culminating in what is now called chaotic dynamics. The ecological difference equation models that Bob studied, together with the differential equations that meteorologist Edward Norton Lorenz had previously analyzed, spurred the mathematical development of chaos theory and placed limits on our ability to make predictions about systems governed by nonlinear dynamics. At his Princeton going-away party in 1988, the celebrants serenaded Bob with a song called "Kid Chaos," composed by his colleague Henry Horn. Bob was quietly proud of this song and enjoyed Henry's reprise at Bob's 70th birthday celebration at the University of Oxford.

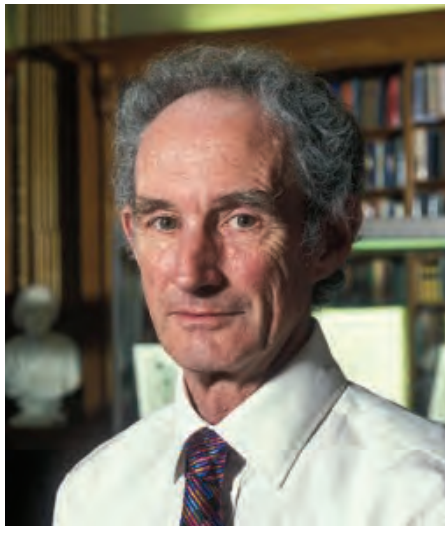


PHOTO: ROB COUSINS/ALAMY STOCK PHOTO

Through most of the 1970s and 1980s, Bob spent his summers at Imperial College London's Silwood Park campus near Windsor, England. There, he initiated a series of long-term collaborations with far-reaching consequences. For example, he and epidemiologist Roy Anderson transformed the field of epidemiological dynamics, and with biologist Michael Hassell he reshaped insect predator-prey dynamics. Summers at Silwood Park were characterized by intense work with a large and changing cast of collaborators, punctuated by equally intense competitive sport in which winning was more important than strict adherence to the rules. The order of authorship for several papers was decided by the outcome of croquet games played on a lawn sculpted by the local rabbits, whose effect on the game only Bob really understood.

After taking a position at the University of Oxford in 1988, Bob quickly became influential in U.K. scientific life and in 1995 was appointed government chief scientific adviser (GCSA). Among Bob's many achievements as GCSA was codifying the role of science advisers, a move that remains critically important today during the coronavirus disease 2019 (COVID-19) pandemic. In this role and later as president of the Royal Society, Bob worked actively, and often behind the scenes, to promote greater inclusivity and encourage early career scientists.

Throughout his life, Bob remained a proud Australian and would often play up to the image of the plain-spoken Aussie. He claimed that his favorite film was the Australian classic *Mad Max*. A British civil servant once tried to summarize Bob's dismissal of a daft idea by saying he "expressed himself in a somewhat…'Australian way'!"

Bob was a master of the witty put-down. With a smile, he once told a very senior botanist, "In my experience, the intellectual capability of any biologist is directly proportional to the speed of locomotion of the organisms with which they work." In scientific debate, he could be cutting, although always with reason. He was immensely kind and loyal to friends and colleagues and enormously encouraging to people at the start of their careers, something we both experienced personally as young academics at Imperial College London and the University of Oxford.

Bob loved being in the wild. For nearly 50 years, he organized an annual hiking trip with colleagues, and he spent most of his holidays walking with his wife, Judith. At the end of one hiking trip, just before the final descent, Bob asked the rest of us to go on ahead so he could have a final moment alone to soak in the mountain ambience. It was this deep attachment to the natural world that we believe motivated his great concern for the future of the environment.

Bob May was knighted, awarded the Order of Merit, and made a peer in the U.K. House of Lords as well as a Companion of the Order of Australia. He received the Crafoord Prize, the Balzan Prize, and many other prizes and honors for his scientific research. He is survived by his wife, Judith, and their daughter, Naomi. Bob leaves the field of ecology stronger and more rigorous than it was when he joined it in the early 1970s, a transformation in great part due to the combined force of his leadership and his intellect. ■

[1]Oxford Martin School, University of Oxford, Oxford, UK.
[2]All Souls College, University of Oxford, Oxford, UK.
Email: charles.godfray@oxfordmartin.ox.ac.uk

10.1126/science.abc7800

POLICY FORUM

RESEARCH POLICY: COVID-19

Moving academic research forward during COVID-19

A gradual, stepwise approach to reopening, informed by public health expertise, will be essential

By **N. S. Wigginton[1,2], R. M. Cunningham[1], R. H. Katz[3], M. E. Lidstrom[4], K. A. Moler[2], D. Wirtz[5], M. T. Zuber[6]**

The coronavirus disease 2019 (COVID-19) pandemic has led to an unprecedented disruption of society. Institutions of higher education have been no exception. To preserve the safety of their communities and adhere to public health guidance, universities and colleges around the world have rapidly pivoted to fully online teaching and learning models, implemented remote work for the majority of employees, and shuttered countless public spaces and programs. Most "on-site" research activities—in laboratories, in clinics, or in the field—also ground to a halt. Many institutions are now planning or implementing a ramp-up of on-site research activities, which offers an opportunity to begin implementing policies and practices that will lay the groundwork for the eventual reopening of additional on-site academic programming, including teaching. To ramp up safely, institutions are working with stakeholder groups—such as public health experts, as well as faculty, staff, and students—to develop guiding principles that will help inform and drive decision-making over the coming months. We synthesized several risk and decision-making frameworks under development at our universities to develop a set of criteria informed by public health expertise that institutions should consider before and during the first stages of restoring research activities and less certain factors to consider for subsequent phases.

Ramping down academic research and development around the world will undoubtedly contribute to the long-term economic ramifications of COVID-19. In addition to supporting the teaching and service missions of higher education—and health care delivery within academic medical centers—academic research contributes greatly to global economic development. In the United States, for example, higher education institutions accounted for $74 billion, or ~13%, of the $580 billion spent nationally on research and development in 2018 (*1*). More critically, these same institutions accounted for nearly half of the $96 billion spent on basic research nationwide, often seen as the seed corn for innovation and industry. Moreover, academic research institutions are among the top five employers in 44 of 50 U.S. states, employing more than 560,000 people (and more than 300,000 trainees) directly on research funds (*2*), many of which cannot perform their work remotely.

Markings are placed on the floor of an empty lab to promote social distancing.

PHOTO: J. FOLLO/UNIVERSITY OF MICHIGAN

RAMPING DOWN

Public health mitigation strategies across the globe have affected on-site research to varying degrees. In China, university research was subject to strict control measures in Wuhan and elsewhere, which contributed to the mitigation of the spread of the virus across the country (*3*). In Australia, where COVID-19 remains under greater control owing to early mitigation efforts, universities moved classes online, but social distancing measures and encouraging nonessential work from home when possible were deemed sufficient to keep most research facilities at least partially open.

In countries and regions where community transmission has been most severe—including the United States, Europe, and China—most academic institutions implemented policies to cease all "nonessential" on-site research activities over a short time frame, in some cases just a few days. This included not only laboratory research in the physical and life sciences but also field-based activities involving travel or direct human contact, such as clinic-based health, social, or educational research. Exemptions for accessing facilities on campus were solely made for work required to maintain equipment, preserve specialized research materials or long-term experiments, perform research to address the ongoing pandemic or other research deemed essential, or ensure patient, animal, and laboratory safety. Although varying widely by discipline and region, we estimate that these restrictions have halted more than 80% of on-site research activity at our six institutions.

RAPID RESPONSE

Despite the myriad challenges associated with ramping down on-site activities, research institutions worked closely with state and federal governments, funders, private industry, and each other to maintain continuity of research operations. In the United States, universities and their associations have been working closely with federal agencies to clarify what activities are allowed under active grants (e.g., salary continuity for researchers who aren't able to work on-site). Other coordination efforts include commitments to open sharing of data and research findings during the pandemic (*4*), improving access to high-performance computing resources for COVID research (*5*), and licensing terms that prioritize access to potentially life-saving technologies (*6*).

Academic researchers have also greatly contributed to work that directly addresses the ongoing pandemic—from revealing the fundamental biology of severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2), to studying the vast social, behavioral, and

[1]University of Michigan, Ann Arbor, MI, USA. [2]Stanford University, Stanford, CA, USA. [3]University of California, Berkeley, CA, USA. [4]University of Washington, Seattle, WA, USA. [5]Johns Hopkins University, Baltimore, MD, USA. [6]Massachusetts Institute of Technology, Cambridge, MA, USA. Email: nwigg@umich.edu

Phased approach and possible mitigations for determining allowable on-site research

	COMMUNITY TRANSMISSION STATUS (*15*)	ON-SITE ACTIVITIES PERMITTED	MITIGATIONS	RESEARCH WORKFORCE IMPLICATIONS
Phase 0 (current state)	Substantial	Only essential work to ensure laboratory, animal, or patient safety; maintenance of equipment, materials, or long-running experiments; COVID-related research	Strict building access; personal protective equipment required; all work done remotely, if possible	Only essential staff with considerable training allowed; travel disrupted; substantial absenteeism owing to illness, child care, or family care; high-risk workers restricted
Phase 1 (ramp-up)	Moderate	Gradual addition of laboratory and studio work and regional field research not involving human subjects; widely used shared facilities reopen	Control building and/or room access; require temperature and symptom checking, physical distancing, strict limits on occupancy in labs, use of masks, enhanced cleaning procedures, and closures of exposed work spaces and buildings; testing and contact tracing if and when available	Fraction of researchers allowed depending on need, training, and willingness to return
Later phases	Minimal to none	Continued gradual addition of more on-site research activities, use of office and shared spaces, and relaxation of travel prohibitions; research with human subjects will require the highest level of scrutiny	Gradual loosening of some control measures, depending on performance metrics	Phased introduction of researchers working in-person with human subjects; additional trainees; high-risk workers only when conditions allow

economic impacts felt across the world, to developing the tests, therapies, and vaccines that will help treat the disease and prevent its transmission. Researchers around the globe have published more than 13,700 papers on SARS-CoV-2 or COVID-19 (*7*) and posted more than 3700 preprints to the bioRxiv and medRxiv repositories as of 19 May 2020.

Institutions are also assisting with critical public health services such as testing and providing the public and decision-makers with real-time data about the pandemic. For example, Johns Hopkins University's international COVID-19 dashboard receives 1.5 billion views each day, providing invaluable data on total confirmed cases, deaths, recovery rates, bed occupancy, intensive care unit availability, and more (*8*). Overall, the collaborations and open sharing of data and knowledge across international borders have proven to be essential in the response to the pandemic and to the reopening of other economic sectors.

GUIDANCE FOR PHASED RAMP-UP

Months after most on-site research was shut down, institutions in China, Europe, and the United States have slowly started resuming on-site research. Institutions have developed principles and policies for resuming on-site research activities based on input from public health and biosecurity experts, faculty, staff, students, and other community members. Our six universities, which represent a range of public and private institutions under varying state and local mandates and levels of local virus transmission, have developed overlapping yet distinct guidance for our research communities (see https://doi.org/10.3886/E119503V1). Common themes within our plans and elsewhere center around the critical need to adhere to public health guidance, prioritize the health and safety of the workforce and participants, and implement fair and transparent processes for decision-making. However, our plans, as well as others around the United States and in other countries, also diverge in ways that may be determined by a host of other factors, from cultural norms on campuses to local and state regulations. Policies such as allowing on-site undergraduate researchers, deciding acceptable occupancy levels in facilities, deciding whether to prioritize certain buildings and activities at the expense of minimal access to everyone, permitting use of nonlaboratory on-site spaces—including libraries, offices, and studios—and allowing field research that does not involve human subjects are among the primary differences in institutional responses. Variations in ramp-down and ramp-up approaches often reflected differing local and state guidelines or mandates, where political and social pressures have the potential to conflict with the best public health recommendations.

Public health frameworks provide a critical and helpful risk-based assessment for when certain industries, governments, and the economy more broadly can reopen [e.g., (*9*)]. Academic institutions represent a broad set of activities and associated risk where one size (and one policy) does not fit all; however, it is clear from public health expertise that a gradual, stepwise approach to reopening and operating will be essential [see the table; (*10*)]. Furthermore, the use of metrics both in the community and within institutions will help determine if and when it is suitable to move into the next phase. On-site testing, contact tracing, and determining immunity status will likely play important roles not just in institutional decision-making and risk mitigation but also for broader public health monitoring (*11*). To do so, institutions will also have to consider how these strategies for research complement strategies being discussed for their broader campus community, as well as weigh potential costs, resources, and privacy concerns. Other metrics that will help determine when institutions are prepared to move into the next phase include building and laboratory occupancy rates, rates of adherence to physical distancing guidelines, and the number of new cases and symptomatic individuals identified during screening (see the box).

Future ramp-up and stabilization phases should be cautious and flexible enough that research activity can also ramp back down if metrics, public health guidance, or other external factors (e.g., local health care system capacity) dictates. Within institutions, this may also be required for certain laboratories, floors, or buildings if cases are identified and researchers are required to self-quarantine after possibly being exposed to a sick co-worker.

Further control measures will be required for months or more, such as continued physical distancing, engineering controls, requiring personal protective equipment, and administrative controls that include staggering access to spaces through shifts to minimize interactions between personnel (*9*). Although our suggestions are intended to prioritize caution and reversibility, we are concerned that other ramp-up plans might instead reopen too quickly or without proper safeguards out of a desire to return to prepandemic operations as soon as possible. As we are seeing in countries or other sectors that are prematurely reopening, undesired outcomes such as new transmission and outbreaks could lead to a whiplash effect of being fully open and then back to fully closed. Gradually and carefully resuming on-site research, and demonstrating that mitigations are effective, provides an ideal opportunity for institutions to implement lessons learned to inform the potential arrival of thousands of undergraduate students when terms resume. It will also help inform when other higher-risk activities, such as in-person work with human subjects, can safely resume.

LOOKING AHEAD

Given the length of time that may be required to continue practicing social distancing, it may be years before academic research institutions reach a new normal. Although some beneficial practices may become more routinized (e.g., more alternative work arrangements and virtual meetings), there will undoubtedly be far more deleterious impacts across higher education. Anticipated budget shortfalls from multiple revenue streams suggest that the ongoing pandemic will hamstring institutions financially for years to come. Regarding research specifically, institutions will have fewer internal resources to perform research, invest in research infrastructure, and maintain its workforce. This presents challenges not only for individual institutions but also for the global research enterprise as a whole. In the United States, for example, institutional investments in research comprised ~25% of total higher education R&D spending in 2018 (*12*), a proportion that has increased considerably over the past decade as the percentage of federal investment in research has declined. For countries in which a large percentage of its research workforce consists of international students, such as Australia, travel and visa restrictions could lead to a substantial loss in revenue to support operations and a considerable reduction of the national scientific workforce (*13*).

The response to COVID-19 has highlighted how the lack of scenario planning and disaster preparedness is a systemic problem spanning virtually all sectors of society. Despite clear guidance and recommendations based on lessons learned from other disasters (*14*), the research community has much work to do to improve disaster resiliency. The experience of COVID-19 should make it clear that resilience planning should be a priority going forward, but even the best laid plans fail without effective leadership and coordination. Global coordinating bodies like the World Health Organization, or national agencies, must not be sidelined in their ability to advise governments and guide policies.

In the absence of strong national leadership, most institutions had to quickly develop their own plans for ramping down research, supplemented by ad hoc communication between institutions. Coordinating bodies like the Association of American Universities, which represents 63 major research universities in the United States and Canada, are playing much more prominent roles in facilitating ramp-up and other long-range planning. Improved coordination across academia, government, health systems, and industry during crises will also help identify early roles that institutions could play to address critical needs. For example, institutions could deploy expertise, resources, or facilities when there is insufficient incentive or capacity for the private or public sectors to refocus production or facilities rapidly, or when they lack capacity to scale up services such as testing. Considering a broader subset of the R&D workforce among essential workers, as in Washington state's "Stay Home, Stay Healthy" order, would help facilitate these cross-sector collaborations more effectively while also maintaining other potential lifesaving research unrelated to the pandemic.

Finally, COVID-19 has exacerbated multiple equity issues in the research enterprise that institutions will grapple with in the months and years ahead. This broad-scale disruption of research operations has led to an incalculable number of setbacks for researchers, many of which disproportionately affect early-career researchers and their career advancement. These include the cancellation of long-running experiments, the loss of opportunities to collect critical data (e.g., in field and clinical studies), and lack of access to specialized major instrumentation, among many others. Furthermore, longstanding affordability and child- and family-care disparities across the research workforce—which disproportionally affect women, lower-income support staff, and trainees—are more clear than ever given the sudden and asynchronous sector closures and cost-saving measures implemented at many institutions. Researchers that fall into higher-risk categories on the basis of preexisting health concerns, age, or other immunocompromising conditions face long-term uncertainties around when it is safe to return to work. Systemic solutions such as extensions to promotion and tenure clocks, further deployment of alternative work arrangements, additional fellowship support for trainees, and policies to allow for extended paid and unpaid leave will be essential to stabilize the research workforce.

Moving forward, it will be up to academic institutions, governments, and funding agencies to develop practices and policies that encourage a more resilient, nimble, and equitable research ecosystem during the COVID-19 pandemic and beyond. Deeper investments in the research workforce and infrastructure will surely help; however, governments should also incentivize stronger ties between public health agencies and academic research institutions to ensure that decision-making at institutions and across communities is guided by the best available research. If not, it is unlikely that the research enterprise or society as a whole will be any better positioned to help generate solutions, or recover itself, when the next disaster arrives. ■

Should we ramp up?

Checklist for assessing when more research activity is permissible on-site

- Does local or regional public health guidance permit a gradual increase in research activity?
- Do individual labs and the institution have reliable access to supplies such as personal protective equipment and disinfectants?
- Does the institution have the ability to track symptoms, conduct testing, and/or trace and inform contacts?
- Have ramp-up procedures and plans been communicated to researchers?
- Are departments and individual labs ready to work safely?
- Are the support units (facilities, environmental health, security, custodial, transportation) prepared to support more on-site activity?

REFERENCES AND NOTES

1. National Center for Science and Engineering Statistics, "National patterns of R&D resources: 2017–18 data update" (NSF 20-307, National Science Foundation, 2019); https://ncses.nsf.gov/pubs/nsf20307
2. The Institute for Research on Innovation & Science, Summary documentation for the IRIS UMETRICS 2019 data release. IRIS (2020); https://doi.org/10.21987/a07t-1c72
3. H. Tian *et al.*, *Science* **368**, 638 (2020).
4. Wellcome, Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak (2020); https://wellcome.ac.uk/coronavirus-covid-19/open-data.
5. The COVID-19 High Performance Computing Consortium; https://covid19-hpc-consortium.org
6. AUTM, COVID-19 licensing guidelines; https://autm.net/about-tech-transfer/covid19/covid-19-licensing-guidelines
7. Q. Chen, A. Allot, Z. Lu, *Nature* **579**, 193 (2020).
8. E. Dong, H. Du, L. Gardner, *Lancet Infect. Dis.* **20**, 533 (2020).
9. C. Rivers *et al.*, "Public health principles for a phased reopening during COVID-19: Guidance for governors" (Johns Hopkins University, 2020); www.centerforhealthsecurity.org/our-work/pubs_archive/pubs-pdfs/2020/reopening-guidance-governors.pdf
10. American College Health Association (ACHA), "Considerations for reopening institutions of higher education in the COVID-19 era" (ACHA, 2020); www.acha.org/documents/resources/guidelines/ACHA_Considerations_for_Reopening_IHEs_in_the_COVID-19_Era_May2020.pdf
11. D. M. Altmann, D. C. Douek, R. J. Boyton, *Lancet* **395**, 1527 (2020).
12. National Center for Science and Engineering Statistics (NCSES), Higher education research and development survey: Fiscal year 2018 (National Science Foundation, 2018); https://ncsesdata.nsf.gov/herd/2018/html/herd18-dt-tab001.html
13. F. Larkins *et al.*, "Impact of the pandemic on Australia's research workforce," 6 May 2020, p. 14; www.science.org.au/sites/default/files/rrif-covid19-research-workforce.pdf
14. National Academies of Sciences, Engineering, and Medicine, "Strengthening the disaster resilience of the academic biomedical research community: Protecting the nation's investment" (Consensus Study Report, National Academies Press, 2017); https://doi.org/10.17226/24827
15. U.S. Centers for Disease Control and Prevention (CDC), "Interim guidance for administrators of US institutions of higher education" (CDC, 2020); www.cdc.gov/coronavirus/2019-ncov/community/guidance-ihe-response.html

ACKNOWLEDGMENTS

We acknowledge helpful discussions and feedback from several individuals at our institutions, in addition to T. Smith, P. Schiffer, J. Walsh, and anonymous reviewers.

Published online 28 May 2020
10.1126/science.abc5599

A biodiversity target based on species extinctions

A single target comparable to the 2°C climate target may help galvanize biodiversity policy

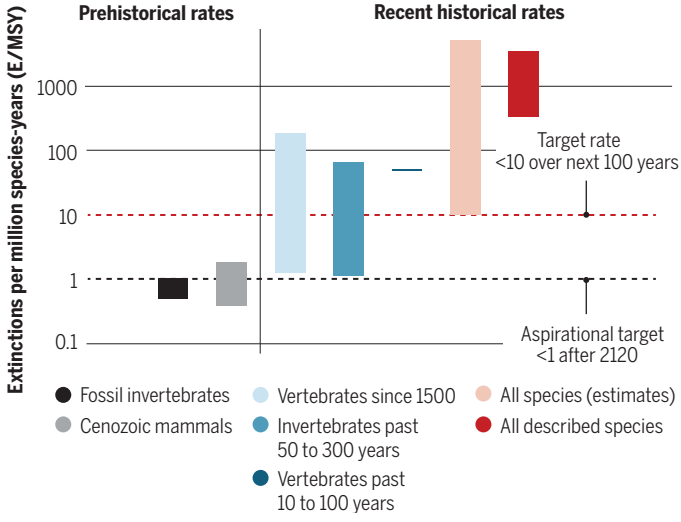
By **Mark D. A. Rounsevell[1,2], Mike Harfoot[3], Paula A. Harrison[4], Tim Newbold[5], Richard D. Gregory[5,6], Georgina M. Mace[5]**

Although worldwide loss of biodiversity arising from human activities is widely known, policy has been unable to arrest the decline (*1*). Much of this failure can be attributed to a lack of mainstreaming of biodiversity in public policy (*2*, pp. 741–762) and limitations in raising the profile of biodiversity loss for politicians and the public. Of the 20 Aichi Biodiversity Targets (ABTs) established in 2010 by the Convention on Biodiversity (CBD), only four show good progress, whereas 12 related to the state of nature show worsening trends (*1*). With the 2020 target date for the ABTs now upon us, it is critical to define a post-2020 agenda to arrest the loss of biodiversity. This will require a target, underpinned by a clear global goal for biodiversity, that can be readily communicated to galvanize both political will and public support. Similarly to how the climate change community uses a single indicator (global mean temperature change) and a target (maximum 2°C rise relative to preindustrial levels) as a rallying point for policy action and agreements, we propose a 2°C-like target for biodiversity (see table S1): a measurable, near-term target of keeping described species extinctions to well below 20 per year over the next 100 years across all major groups (fungi, plants, invertebrates, and vertebrates) and across all ecosystem types (marine, freshwater, and terrestrial).

Although there are many comprehensive proposals that seek to contribute to the post-2020 agenda (*3*, *4*), they focus on achieving conservation actions, such as increasing the coverage of areas dedicated to wildlife, or maintaining intact wilderness, rather than specifying required outcomes for biodiversity (*5*). As such, there is a real risk, as with some of the 2020 ABTs, that targets will be met, yet biodiversity will continue to decline (*5*–*7*). We propose an alternative approach that draws on the theory of change (see tables S1 to S3). We suggest that the overall impact of a post-2020 biodiversity framework should be the goal of achieving the CBD 2050 vision of "Living in harmony with nature." However, measuring progress toward this goal requires a communicable and actionable indicator and target, which for biodiversity has proven challenging.

Targeting an extinction rate

Extinction rates (E/MSY) across a variety of taxonomic groups for different historical periods are related to the proposed extinction rate target for the next 100 years and the aspirational target (background extinction rates) from 2120. Bars show the full range of possible values for E/MSY when E, S, and Y are represented by ranges of possible values (see Table S4 for the data sources). Data encompass all plants, animals, and fungi unless indicated otherwise.



As a largely political construct, supported by science, the 2°C climate policy target adopted by the United Nations Framework Convention on Climate Change (UNFCCC) was not intended to represent the multiple dimensions of the climate system and the diverse impacts of climate change. The target has the great advantage of being communicable to a nonscience audience and so supports understanding of the ambition in limiting climate change. Similarly, we argue that a comparable simple and measurable indicator is needed to support biodiversity policy with a specific and easily communicated target against which policy responses can be developed and tested. As with the global mean temperature metric for climate change, a single biodiversity metric will inevitably mask considerable spatial variation in the status of biodiversity and ignore many of the complexities inherent in ecological systems. However, as with climate change, a communicable, 2°C-like target for biodiversity should be supported by a broader range of indicators and targets that more fully describe the state of biodiversity and its drivers of change.

SPECIES EXTINCTION–BASED TARGETS

A single biodiversity metric needs to be both relevant and effective for the goal, as well as being easy to measure and communicate. With this in mind, we propose a metric based on the rate of species extinction. Current species extinction rates clearly exceed those that were characteristic of the past, and projected future extinction rates are much higher still, even given large uncertainties (*8*). Over the coming decades, some continuing loss of species is inevitable given the current human domination of Earth's systems, so we suggest an ambitious but achievable rate: keeping described species extinctions to well below 20 per year over the next 100 years (see the figure). Thereafter, once there is stabilization of human impacts, we suggest that a rate closer to background rates (i.e., prehistorical rates) should be the aspiration (see the figure).

The figure of 20 extinctions per year is based on a target to reduce extinctions to 10 per million-species-years applied to the estimated 2 million species described by science (see the figure and the supplementary

[1]Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU), Karlsruhe Institute of Technology, Garmisch-Partenkirchen 82467, Germany. [2]School of GeoSciences, University of Edinburgh, Edinburgh EH8 9XP, UK. [3]UN Environment World Conservation Monitoring Centre (UNEP-WCMC), Cambridge CB3 0DL, UK. [4]UK Centre for Ecology and Hydrology, Lancaster, LA1 4AP, UK. [5]Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK. [6]RSPB Centre for Conservation Science, the Lodge, Sandy, Bedfordshire SG19 2DL, UK. Email: mark.rounsevell@kit.edu

GRAPHIC: X. LIU/*SCIENCE*

materials). The 10 per million-species-years is the threshold adopted in the planetary boundaries framework (*9*). This threshold was used to reflect the addition of a large uncertainty bound to the upper estimate of natural background extinction rates of 1 per million-species-years (*9*). In proposing this single metric, we are not suggesting that it is sufficient on its own to describe the changing state of biodiversity, or to guide conservation policy, or ecosystem management on the ground. However, we suggest that the extinction rate is a necessary element of any biodiversity policy target and that it embraces the core concerns that most people have about biodiversity loss.

Our justification for basing a global target only on the rate of species extinction is twofold. First, extinction fully incorporates the most fundamental aspect of biodiversity loss. Each species embodies distinct genetic diversity, usually shaped and developed over millions of years of independent evolution. The total body of extant species is the diversity of life on Earth. The extinction of a species represents an irreversible loss, a measurable reduction in the diversity of life on Earth, and is the ultimate concern for conservation. Although changes in species abundance or to ecological communities may be of equivalent concern (*10*), they are in principle, at least, reversible and recoverable. The extinction rate incorporates loss of species and of genetic diversity and therefore two of the core components of the CBD definition of biodiversity.

Second, species extinction is widely understood and easy to communicate. There is widespread public concern about extinctions, as was demonstrated recently by the emphasis on the number of species at risk of extinction in the media coverage of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) Global Assessment (*1*). There are collaborative networks recording the world's known and extinct species (www.gbif.org/what-is-gbif and www.catalogueoflife.org) and a program of work to catalog species close to extinction and to monitor extinctions as they take place (www.iucnredlist.org). We anticipate that targets such as we propose would stimulate new data and innovative approaches to monitoring and modeling the status of the world's species.

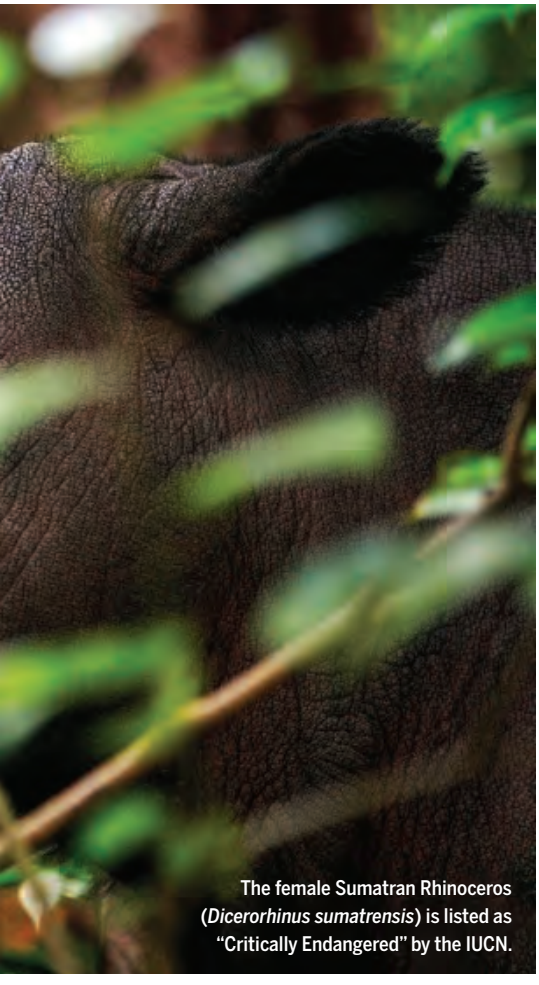
We recognize that measuring only the extinction rate can miss many changes to biodiversity that are of great importance to people and that underpin the sustainable flow of ecosystem services. It would be possible, in principle, to meet a simple target to reduce the rate of species extinction and yet see wholesale and damaging changes to life on Earth. At the extreme, a single, small population of each species, maintained somewhere (perhaps in a protected area), would result in no extinctions, yet could place ecosystem functioning at risk everywhere and fail to meet most people's vision for life on Earth. Achieving our target could therefore have some perverse solutions, just as achieving the 2°C target can be achieved through actions such as solar radiation management, which treat symptoms, but do not address the root problem. Implementation, therefore, requires effective policy scrutiny. Current extinction events are, in any case, recorded [e.g., the annotation of "possibly extinct" by the International Union for Conservation of Nature (IUCN)], and this can contribute to more precautionary reporting.

PHOTO: CYRIL RUOSO/MINDEN PICTURES

POLICY, IMPLEMENTATION, SCALING

In addition to the global extinction rate, additional targets will be needed to ensure that biodiversity meets functional and cultural roles that are especially relevant at local and national scales. This will require countries to develop national targets that are relevant to their own circumstances. For climate change, national targets are defined by Nationally Determined Contributions (NDCs), which outline governmental action toward overall climate mitigation. Likewise, Nationally Determined Contributions for Biodiversity (NDCBs) could provide an action-based context for better protection and management of biological resources and provide governments with a framework within which they could act to achieve global biodiversity targets, including to reduce extinction rates. Governments also need to monitor the effects of NDCBs in real time and use models that project into the future, to evaluate compliance with policy implementation and fine-tune their policy responses.

We are not suggesting that the species extinctions target would be allocated to countries in a top-down policy process. NDCs are voluntary, but monitored through the (independent) Climate Action Tracker to check if countries' national targets (which are decided by each country) collectively achieve the global target. One could envisage something similar for biodiversity where country NDCBs for actions within their own country boundaries and actions related to teleconnections are related to both national and global targets.



The female Sumatran Rhinoceros (*Dicerorhinus sumatrensis*) is listed as "Critically Endangered" by the IUCN.

This might lead to a "Biodiversity Action Tracker" or something similar. Then, as with the UNFCCC, an iterative review process would encourage countries to increase their ambition if collectively the global targets were not being achieved.

Generally, countries with the most critically endangered species have the best opportunity to take direct action to mitigate the direct drivers pushing them toward extinction. Hence, national extinction targets need to be clear and sum up to the shared global ambition. Because the proposed global target concerns the number of species becoming extinct within 100 years, it follows that a national target will need to refer to national responsibilities toward reducing the number of global extinctions. National responsibility refers to the geographic location where removal of direct pressures on a species needs to occur, e.g., responsibility for reducing the rate of habitat loss for species in the Atlantic Forest of Brazil would lie in Brazil.

Many other approaches, including the ABTs, also focus at the national level. We feel that it is important to highlight the need for national responsibility for the global target. In a world increasingly interconnected by trade and the impacts of consumption, the drivers of biodiversity loss often originate, at least in part, in other countries because of consumer demand (*11*). It is important, therefore, for nations to recognize the displacement effects of consumption patterns on biodiversity, along with the moral responsibility to act on this problem. Hence, it is equally important for nations to commit to the direct conservation of species beyond their national borders, because the risk of species extinctions is not equally distributed geographically. A commitment to support biodiversity in other parts of the world would mirror (and complement) efforts to alleviate poverty through international actions that foster development and would tackle the asymmetries inherent in financially poor countries often being biodiversity rich. Thus, the national responsibility for the recovery of species may best be addressed with financial aid, trade agreements, and other means to reduce indirect pressures.

WAYS FORWARD

A refreshed vision that lays out an achievable ambition to reverse the decline in biodiversity is urgent. The time is right; there is currently a set of important preconditions to underpin successful policy development and implementation. Examples include the policy-relevant information from IPBES assessments, the ambitious goals embodied in the UN Sustainable Development Goals (SDGs) for 2030, and an emerging consensus that links biodiversity recovery to climate change mitigation and adaptation, and to sustainable development and human well-being, as well as the increasing public awareness and demand for action. Together these provide a valuable opportunity and one for which the world has all the necessary actors and institutions to exploit.

There is a lot of debate about how effective the NDCs have been (which is relevant to assessing how effective NDCBs might be). However, in the end, the effectiveness of outcomes is strongly dependent on governmental willingness to act. This affects any indicator. The advantage of having a single, simple indicator as we propose here is that it makes lack of achievement toward the target more visible. It also makes clear that national governments have responsibility for reducing extinction rates everywhere.

There is an important role for science in supporting this policy process by advancing understanding of some of the key elements of the theory of change. We need better understanding of the impact of policy actions on biodiversity outcomes and how this supports policy goals, e.g., through a Biodiversity Action Tracker. We need better modeling and futures analyses (target-seeking and ex-ante scenarios) to understand how sets of time-dependent actions encapsulated in alternative development pathways can move us toward the vision (*2*). We also need scientific advances to investigate what levels of biodiversity loss are damaging for ecosystem functioning, and to fill knowledge gaps about the complex interactions between species extinctions, biodiversity, ecosystem functioning, and ecosystem services.

We recognize that a single indicator cannot capture the multiple dimensions of complex systems, and this holds as much for biodiversity assessment as it does for climate change. Ultimately, however, we need to catalyze both policy and public support for biodiversity and its preservation, and to do this we need an indicator and target that readily communicate the urgency of the problem to multiple audiences and enable monitoring of progress. Our proposal for a 2°C-like target for biodiversity based on the rate of global species extinctions attempts to provide a simple, measurable, and easily communicated target against which policy responses can be established and tested. Adoption of such a target globally, and its elaboration and implementation nationally, would help integrate the biodiversity and climate change agendas, which is critical for protecting and managing nature sustainably in achieving the SDGs. ■

REFERENCES AND NOTES

1. IPBES, Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, E. S. Brondizio, J. Settele, S. Díaz, H. T. Ngo, Eds. (IPBES Secretariat, Bonn, Germany, 2019).
2. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, The Regional Assessment Report on Biodiversity and Ecosystem Services for Europe and Central Asia, M. D. A. Rounsevell, M. Fischer, A. Torre-Marin Rando, A. Mader, Eds. (IPBES Secretariat, Bonn, Germany, 2018).
3. E. Dinerstein *et al.*, *Sci. Adv.* **5**, eaaw2869 (2019).
4. J. E. Watson *et al.*, *Nature* **563**, 27 (2018).
5. P. Visconti *et al.*, *Science* **364**, 239 (2019).
6. L. Boitani *et al.*, *PLOS Biol.* **6**, e66 (2008).
7. D. P. Tittensor *et al.*, *Science* **346**, 241 (2014).
8. V. Proença, H. Pereira, "Comparing extinction rates: Past, present, and future" in Encyclopaedia of Biodiversity (Elsevier, 2017).
9. W. Steffen *et al.*, *Science* **347**, 1259855 (2015).
10. G. M. Mace *et al.*, *Nat. Sustain.* **1**, 448 (2018).
11. A. Marques *et al.*, *Nat. Ecol. Evol.* **3**, 628 (2019).

ACKNOWLEDGMENTS

We thank A. Arneth, N. Burgess, S. Butchart, A. Purvis, P. Visconti, and J. Watson for comments. M.D.A.R. acknowledges financial support of the Helmholtz Association. T.N. acknowledges funding from the Royal Society. The authors declare no competing interests.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1193/supp/DC1

10.1126/science.aba6592

Crowds flocked to Venice Beach and other popular U.S. tourist destinations on Memorial Day.

PHOTO: ROGER KISBY/REDUX

ECONOMICS

Discounting lives

Are COVID-19 deaths "sacrifices"? If so, to what?

By **Audra J. Wolfe**

In late March of this year, Texas Lieutenant Governor Dan Patrick suggested in an interview that many people over 70—himself included—would be willing to risk contracting coronavirus disease 2019 (COVID-19) so as not to, in his words, "sacrifice the country." At the time, his comments were widely reviled. Just over 2 months later, the public appears to have accepted the calculus that lives will in fact have to be sacrificed in the name of the economy.

Emerging reports make clear that the burden of this "sacrifice" is not being borne equally. As of mid-April, in Richmond, Virginia, a city that is 48% Black, every single person who had died from the disease was Black (*1*). Similarly, the population of Chicago is ~30% Black, but, as of 28 April, Black people made up 54% of the city's COVID-19 deaths (*2*). The *New York Times* roster of COVID-19 "cluster" sites nearly exclusively lists prisons, jails, meat processing plants, and nursing homes (*3*).

In the United States, the lives of people of color and those of the poor, the elderly, the disabled, and the incarcerated are heavily discounted in the economic sense, a phenomenon documented in Howard Steven Friedman's *Ultimate Price*. Written for a lay audience, Friedman's book explains how the U.S. government and corporations assign dollar values to human lives. Chapters survey the techniques decision-makers use in leveling penalties for wrongful death, evaluating potential regulations, assessing corporate liability, setting life insurance premiums, providing health care, and choosing to have a child or go to war.

"Price tags" for human beings, Friedman repeatedly shows, reflect existing inequalities in U.S. society. The formula that the government used to compensate the families of victims of the September 11 terrorist attacks, for example, factored in the deceased's annual salary, which meant that the lives of bond traders were valued many times higher than those of restaurant staff.

Early and often, Friedman argues that we must devise more equitable ways to assign value to human life. He is not wrong. Even so, the book has reformist rather than radical goals: Readers are exhorted to understand how lives are priced so that they might demand better formulas.

As a counterpoint to Friedman's grim pragmatism, I found it instructive to read Michelle Murphy's 2017 book *The Economization of Life* alongside *Ultimate Price*. Murphy, like Friedman, is intensely interested in how experts have assigned value to human populations. But where Friedman treats economics as a solid basis for rational decision-making, if done right, Murphy explores how the concepts of "economy" and "population" gained their appeal.

Through a series of chapters that explore the history of ideas about population control in the United States and Bangladesh in the mid-to-late 20th century, Murphy persuasively argues that life has come to be valued primarily in terms of its ability to contribute to the macroeconomy of nation-states. This has not come to pass through the action of free markets but rather through social scientific practices, including indexing, surveying, and other forms of counting, that re-create racial hierarchies even while rejecting the language of race.

In this moment when we are being asked to weigh the risks of (some people's) deaths against the economic costs of stay-at-home orders, Murphy's notion of the economy as a "phantasmagram" is compelling. She uses the term to capture the emotion and aspiration associated with the disembodied metrics we use to measure and track the economy. As phantasmagrams, Murphy writes, quantitative measures such as gross domestic product "have supernatural effects in surplus of their rational precepts." They "conjure ineffable realms that can take shape as a collective phantasy in excess of the representational and logical limits of quantification practices themselves."

On Fridays, the U.S. government releases data on new unemployment insurance claims. Just as Murphy suggests, the weekly release of these numbers stimulates anxiety and prognostication mostly disconnected from the (noneconomic) value of an individual life. The shock value of these numbers pushes the other shocking number, the cumulative number of COVID-19 deaths in the United States—108,194 as of 5 June 2020 (*3*)—to the side. ■

Ultimate Price: The Value We Place on Life
Howard Steven Friedman
University of California Press, 2020. 232 pp.

The Economization of Life
Michelle Murphy
Duke University Press, 2017. 232 pp.

The reviewer is a Philadelphia-based writer and editor and the author of *Freedom's Laboratory: The Cold War Struggle for the Soul of Science* (Johns Hopkins University Press, 2018). Email: audrajwolfe@gmail.com

REFERENCES AND NOTES

1. A. Rodriguez Espinoza, "African Americans make up all of Richmond Coronavirus deaths," VPM News (15 April 2020).
2. S. Schering, "Data provides racial demographics of those who died from COVID-19 in Cook County and suburbs," *Chicago Tribune* (28 April 2020).
3. Coronavirus in the U.S.: Latest Map and Case Count, *The New York Times*; www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html#clusters.

10.1126/science.abc0164

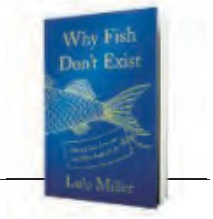
SCIENCE LIVES

Labeling fish, labeling people

A science reporter grapples with an ichthyologist's all-consuming passion for categorization

By **Elizabeth Landau**

Why Fish Don't Exist: A Story of Loss, Love, and the Hidden Order of Life
Lulu Miller
Simon and Schuster, 2020. 240 pp.



Part of being human is labeling things. We mentally label people, animals, and objects, sorting them into separate categories in our minds. When we encounter unfamiliar entities, we attempt to match what we see to our cerebral collections. But sometimes our internal labeling system is based on fundamental misconceptions, and this can have disastrous consequences.

Why Fish Don't Exist by Lulu Miller tells the story of David Starr Jordan, who, together with his students, discovered more than 2500 fish species at the turn of the 20th century—nearly 20% of the fish known at the time. As Miller reveals, however, his ironclad belief in categorization and the presence of a hierarchy in nature ultimately led him to promote misguided and morally abhorrent social policies.

Like its take-home message, this extraordinary book defies classification. It is a memoir, a biography, and possibly a murder mystery. Miller, a gifted storyteller known for her work on the popular NPR podcast *Invisibilia*, uses her own philosophical crises and curiosities to drive the narrative forward and grants the reader a front-row seat to her emotional and intellectual journey.

The book's curious title refers to the idea that many taxonomists no longer consider "fish" a meaningful category. A lungfish, for example, because of its lunglike organs and the structure of its heart, is "more closely related to the cow than to the salmon," Miller writes. Yet we continue to use "fish" as a catchall label for many water-dwelling species. This, she hypothesizes, formalizes their status as "other" and helps us normalize our view of these species as food and as objects to be exploited.

Born in 1851, Jordan marveled at the natural world from a young age. After graduating from Cornell University, he attended a summer course in natural history on Penikese Island, off the coast of Massachusetts, that had been established by prominent biologist Louis Agassiz. "Here I made my first acquaintance with fishes of the sea," Jordan wrote in the first volume of his memoir, *The Days of a Man*. Energized by the experience and "his prophet Louis Agassiz," Jordan set out to make a career of naming fish. He would become the foremost ichthyologist of his time and, at age 40, was appointed Stanford University's first president.

Miller first heard of Jordan on a tour of a science museum in California, where she learned how he had bounced back from near-ruin after an earthquake almost destroyed 30 years of his life's work. Years later, devastated by a breakup, she remembered this anecdote and decided to look deeper into the scientist's life in the hope that his story might offer useful lessons in how to persevere in the face of adversity.

Valenciennea violifera **(top) and** ***Dendrochirus sausaulele*** **(bottom)**

Jordan, it turned out, had overcome not one but many major setbacks. When a fire destroyed his collections in 1883—"[t]he jars would have exploded like tiny bombs," notes Miller—he did not back down from his quest to order nature. And when, on 18 April 1906, an earthquake ruined many of his specimens and separated others from their labeled jars, the taxonomist set about sewing tin name tags "directly to the flesh" of the specimens he recognized, a practice he would continue for the rest of his career. "It was a small innovation with a defiant wish, that his work would now be protected against the onslaughts of Chaos, that his order would stand tall next time she struck," Miller writes.

But Jordan also became engrossed in a different labeling scheme, one promoting a eugenics movement that led to the forcible sterilization of tens of thousands of Americans. In an article published in 1898 and in several later books, he advocated "for the cleansing of the gene pool," Miller writes. His efforts helped pass the first eugenics sterilization laws in the United States in the early 1900s. The consequences of those laws reverberate today. In a particularly heartbreaking chapter, Miller meets a woman named Anna who, in 1967 at the age of 19, had been sterilized against her will. "I always wanted kids, but I couldn't have none," Anna tells her.

The term "eugenics," Miller notes, was coined by Charles Darwin's half-cousin Francis Galton, who, like Jordan, failed to appreciate Darwin's assertion that variation in genes, and in species, leads to evolutionary health and success. "The fact that many of his fish were in fact discovered by the very targets of his eugenicist campaign—the immigrants and 'paupers' whose value to society he dismissed—was something David chose to omit from the scientific record," she observes.

Jordan went to his grave believing in hierarchies, labels, and categories. But Miller is left skeptical of this worldview. In the personal thread of the book, she describes finding happiness with someone who differs from her initial expectations for a long-term romantic partner. The best fit, she surmises, often lies outside of our predefined boxes. ■

The reviewer is a science writer in Washington, DC 20036, USA. Twitter: @lizlandau

10.1126/science.abc6653

IMAGES: "THE FISHES OF SAMOA," DAVID STAR JORDAN & ALVIN SEALE/SMITHSONIAN LIBRARIES/INTERNET ARCHIVE

The fragile Sundarbans mangrove ecosystem (dark green in this satellite image) is threatened by increased development in Bangladesh.

Edited by **Jennifer Sills**

Saving the Sundarbans from development

The Padma Bridge construction project in Bangladesh is near completion (*1*). The 6.15-km-long bridge, costing nearly US$4 billion (*2*), will connect the northern and eastern parts of the country to the southwest Khulna division, which includes the Sundarbans—the world's largest mangrove forest stretching over 10,000 km^2 in Bangladesh and India (*3*). The bridge, when commissioned, is expected to boost Bangladesh's gross domestic product by as much as 1.2% (*4*), but it will also put the fragile Sundarbans mangrove ecosystem at risk.

Bangladesh has already lost the Chakaria Sundarbans, one of the oldest mangrove forests in South and Southeast Asia, as a result of the rise of commercially profitable shrimp farming (*5*). The same could happen in the Khulna division, where the bridge construction has exacerbated increasing land prices and expanding development of real estate, fisheries, tourism, and resorts near the Sundarbans (*6*). The bridge will also facilitate greater access to the busy Mongla Port and likely accelerate forest clearing and land grabbing in the area.

Although economic gains are important, ecological and environmental issues must be considered in long-term development plans. The Sundarbans mangroves region is a UN Educational, Scientific, and Cultural Organization World Heritage Site, a Ramsar site, and a Class 3 tiger conservation landscape of global priority (*7*). At least 355 species of birds, 49 species of mammals (including the globally endangered Bengal tiger), 87 species of reptiles, 14 amphibians, 291 species of fish, and 334 species of plants have been recorded in the area (*3*). The Sundarbans region is situated on the highly active Ganges-Brahmaputra delta, the third largest river basin in the world after the Amazon and the Congo (*8*). One-third of the Sundarbans consists of rivers and streams, which act as a nursery for fish and other aquatic life, including Asia's last two remaining freshwater dolphin species: the Ganges river dolphin and Irrawaddy dolphin (*9*).

More than 3.5 million people depend on the Sundarbans ecosystem for their livelihood and income (*10*). The world-renowned mangrove forest regularly protects human lives and habitation from cyclones and tidal surges (*11*). The Sundarbans region also plays an important role in regulating key ecological processes, including carbon sequestration, storage, and cycling (*12*).

The Sundarbans forest already faces many challenges, including sea-level rise, salinity intrusion, habitat degradation, and biodiversity loss (*3*). The government's plan to establish a coal-based power plant has also threatened the region's sustainability (*11*). Once traffic begins entering Khulna by way of the Padma Bridge, the risks will only increase. The Government of Bangladesh should provide clear guidelines for development, including strict regulations on land clearing, landfilling, waste disposal, and wastewater discharge. To protect the Sundarbans, any development in the region should also comply with policies that conserve local ecosystems and livelihoods.

Sharif A. Mukul[1,2]*, Saleemul Huq[2,3], John Herbohn[1,4], Nathalie Seddon[5], William F. Laurance[6]
[1]Tropical Forests and People Research Centre, University of the Sunshine Coast, Maroochydore DC, QLD 4558, Australia. [2]International Centre for Climate Change and Development, Dhaka 1229, Bangladesh. [3]International Institute for Environment and Development, London WC1X 8NH, UK. [4]Sustainable Minerals Institute, The University of Queensland, Brisbane, QLD 4072, Australia. [5]Nature-based Solutions Initiative, Department of Zoology, University of Oxford, Oxford OX1 3SZ, UK. [6]Centre for Tropical Environmental and Sustainability Science, James Cook University, Cairns, QLD 4878, Australia.
*Corresponding author. Email: smukul@usc.edu.au

REFERENCES AND NOTES

1. Government of Bangladesh, "Padma multipurpose bridge project" (2015); www.padmabridge.gov.bd.
2. J. J. Hamre *et al.*, "Padma bridge (construction)" (Reconnecting Asia, 2020).
3. S. A. Mukul *et al.*, *Sci. Tot. Environ.* **663**, 830 (2019).
4. J. Islam, "Padma Bridge will up GDP by 1.2 pc," *The Daily Observer* (2015).
5. S. R. Biswas *et al.*, *Wetlands Ecol. Manage.* **17**, 365 (2009).
6. M. Anam *et al.*, "Padma bridge–new lifeline of development," *The Daily Star* (2016).
7. M. M. H. Khan *et al.*, "Bangladesh tiger action plan, 2018–2017" (Bangladesh Forest Department, Dhaka, 2018).
8. G. Rasul, *Int. J. River Basin Manage.* **13**, 387 (2015).
9. B. D. Smith *et al.*, *Mar. Mammal Sci.* **22**, 527 (2006).
10. M. S. Uddin *et al.*, *Ecosyst. Serv.* **5**, e88 (2013).
11. P. Schwartzstein, "This vanishing forest protects the coasts—and lives—of two countries," *National Geographic* (2019).
12. D. C. Donato *et al.*, *Nat. Geosci.* **4**, 293 (2011).

10.1126/science.abb9448

Aggregating data from COVID-19 trials

In their Policy Forum "A strategic approach to COVID-19 vaccine R&D" (29 May, p. 948), L. Corey *et al.* discuss the importance of coordinating randomized clinical trial (RCT) protocols to facilitate the evaluation of coronavirus

PHOTO: NASA, CREATED BY JESSE ALLEN, EARTH OBSERVATORY, USING DATA OBTAINED FROM THE UNIVERSITY OF MARYLAND'S GLOBAL LAND COVER FACILITY

disease 2019 (COVID-19) vaccines, and they highlight the ACTIV (Accelerating COVID-19 Therapeutic Interventions and Vaccines) public-private partnership as one example of productive collaboration. We agree that coordination across RCTs is crucial to ensure that evidence for the treatment and prevention of COVID-19 is adjudicated and disseminated as quickly and reliably as possible. In the absence of coordination, false positives from underpowered and uncoordinated collections of redundant trials could fuel the proliferation of ineffective and potentially dangerous treatments. ACTIV provides infrastructure to coordinate efforts by pharmaceutical companies developing vaccines and novel compounds; a similar platform is needed for voluntary collaboration by diverse partners on the full spectrum of research questions.

To increase the power of RCTs, we have created a pilot repository for RCT protocols led by principal investigators who are open to various levels of collaboration. On the COVID-19 Collaboration Platform (CovidCP), researchers can submit their draft or completed protocols and find collaborators. Together, they can initiate new multi-site trials, work to create collaborative protocols that can be used at multiple sites but as independent studies, admit new research sites under the existing trial and Institutional Review Board, share anonymized interim and final data with sites that choose to conduct a trial under a similar but not identical protocol, and collaborate on data collection tools, data standards, and case report forms.

Organizing multi-site RCTs and, where that is not possible, combining data from separate but similar trials with the use of appropriate subject-level or meta-analytical methods, will produce answers faster and more accurately than conducting each trial independently. Every patient participating in an RCT has the right to have their data used as efficiently and meaningfully as possible. Streamlining protocols can help researchers make full use of data even from trials that are stopped early as a result of a change in standard of care or local epidemic waning, from trials that are small and possibly underpowered, and from single-arm trials.

We trust that the clinical research community will share their work and knowledge in the service of developing all possible tools for fighting this pandemic, and we invite additional input and partnerships to maximize the effectiveness of CovidCP. By working together to determine how a cooperative platform such as CovidCP can most benefit researchers, clinicians, policy-makers, and patients, we can address the COVID-19 pandemic and prepare for future global health emergencies.

Elizabeth L. Ogburn[1]*, Barbara E. Bierer[2,3], Ron Brookmeyer[4], Christine Choirat[5], Natalie E. Dean[6], Victor De Gruttola[7], Susan S. Ellenberg[8], M. Elizabeth Halloran[9,10], Daniel F. Hanley, Jr.[11], Joseph K. Lee[12], Rui Wang[3,13], Daniel O. Scharfstein[1]
[1]Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205, USA. [2]Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA. [3]Harvard Medical School, Boston, MA 02115, USA. [4]Fielding School of Public Health and Department of Biostatistics, University of California Los Angeles, Los Angeles, CA 90095, USA. [5]Swiss Data Science Center, ETH Zürich and EPFL, 1015 Lausanne, Switzerland. [6]Department of Biostatistics, University of Florida, Gainesville, FL 32611, USA. [7]Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, MA 02115, USA. [8]Department of Biostatistics, Epidemiology, and Informatics and Department of Medical Ethics and Health Policy, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. [9]Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA. [10]Department of Biostatistics, University of Washington, Seattle, WA 98195, USA. [11]Johns Hopkins Institute of Clinical and Translational Research, Baltimore, MD 21202, USA. [12]Covid-19 Collaboration Platform, Boston, MA 02118, USA. [13]Department of Population Medicine, Harvard Pilgrim Health Care Institute, Boston, MA 02215, USA.
*Corresponding author. Email: eogburn@jhsph.edu

COMPETING INTERESTS

B.E.B., E.L.O., and D.O.S. are members of the CovidCP executive committee. R.B., C.C., N.E.D., V.D.G., S.S.E., M.E.H., D.F.H. Jr., J.K.L., and R.W. are members of the CovidCP advisory board. S.S.E. serves on the Research Oversight Committee for COVID-19 trials conducted at the Perelman School of Medicine, University of Pennsylvania, and serves on a data monitoring committee for two trials of COVID-19 treatments being conducted at Stanford University.

10.1126/science.abc8993

Brazil undermines parks by relocating staff

In May, in the midst of the coronavirus disease 2019 (COVID-19) pandemic, Brazil's Ministry of Environment issued a series of new acts dismissing and relocating the heads of National Parks administrated by the Federal Agency for Conservation of Biodiversity (*1*). The new management model will relocate experienced technical staff of all National Parks and Reserves into a few Management Centers (NGIs), often far from the Protected Areas (PAs) they oversee. This remote management strategy will result in a substantial reduction of the country's ability to preserve key ecosystems, such as those located in the Brazilian Atlantic Forest of northern Rio de Janeiro state, where the last populations of the highly threatened Golden Lion tamarin (*Leontopithecus rosalia*) live, along with many other endangered species (*2*, *3*).

NGIs were originally designed to integrate and strengthen similar PAs, but few meet this criterion. The Golden Lion tamarins live in tropical forests, but the designated NGI for that region will also include the Restinga de Jurubatiba National Park (*4*), formed by coastal lagoons and sandy plain shrublands. Long-term ecological studies that have been conducted in this park for decades (*5*) are now on hold because of uncertainties posed by the implementation of NGIs. This park is located in one of the most populated regions of South America (*6*), and proper conservation of its unique biota is also in severe peril. Illegal hunting and fishing are common in all Atlantic Forest PAs, despite the presence of PA staff. Removing the staff from the local region will impair effective responses to these activities, likely increasing the loss of rare and threatened species of birds and mammals.

The Brazilian Atlantic Forest is also a potential hotspot for emerging diseases (*7*) given the severe biodiversity impacts of rampant deforestation (*8*). Instead of dismantling the management structure that protects PAs, Brazil should increase the technical and financial support needed by its parks to foster tourism, education, science, and conservation.

Pablo R. Gonçalves[1]*, Fabio Di Dario[1], Ana C. Petry[1], Rodrigo L. Martins[1], Rodrigo N. da Fonseca[1], Malinda D. Henry[1], Francisco de Assis Esteves[1], Carlos R. Ruiz-Miranda[2,3], Leandro R. Monteiro[2], Marcelo T. Nascimento[2]
[1]Instituto de Biodiversidade e Sustentabilidade – NUPEM/UFRJ, Universidade Federal do Rio de Janeiro, Macaé, RJ 27965-045, Brazil. [2]Laboratório de Ciências Ambientais, Universidade Estadual do Norte Fluminense, Campos dos Goytacazes, RJ 28013-602, Brazil. [3]Associação Mico-Leão-Dourado, Silva Jardim, RJ 28820-000, Brazil.
*Corresponding author. Email: prg@acd.ufrj.br

REFERENCES AND NOTES

1. MMA/ICMBio, Portarias 362 a 405 de 8 de maio de 2020, Diário Oficial da União (2020), vol. 89, pp. 30–31 [in Portuguese].
2. J. M. Dietz *et al.*, *Sci. Rep.* **9**, 12926 (2019).
3. C. N. Jenkins *et al.*, *Nat. Conserv.* **9**, 152 (2011).
4. Procuradoria da República no Rio de Janeiro, "MPF recomenda ao ICMBio que revogue as exonerações nas Unidades de Conservação do Norte Fluminense e Região dos Lagos" (Ministério Público Federal, 2020) [in Portuguese].
5. F. de A. Esteves *et al.*, "Restingas and Coastal Lagoons in Northern Rio de Janeiro State: Brazilian Long-Term Ecological Research" (PELD-RLaC, 2020); https://peld.macae.ufrj.br/ [in Portuguese].
6. J. N. Williams, in *Biodiversity Hotspots*, F. E. Zachos, J. C. Habel, Eds. (Springer Berlin Heidelberg, 2011), vol. 53, pp. 61–81.
7. T. Allen *et al.*, *Nat. Commun.* **8**, 1 (2017).
8. J. Settele *et al.*, "COVID-19 stimulus measures must save lives, protect livelihoods, and safeguard nature to reduce the risk of future pandemics" (IPBES, 2020).

10.1126/science.abc8297

Electrophysiology Systems

Bundled Configurations

- Proven designs in a single system
- Discounted pricing with savings and value
- Bundled with 2 manipulators
- Easy toggle selection of active components
- Manipulators and motorized components can be controlled by a single ROE input device
- All component features retained
- Platform systems include rotating bases
- USB interface

SUTTER INSTRUMENT

PHONE: +1.415.883.0128
FAX: +1.415.883.0572
EMAIL: INFO@SUTTER.COM
WWW.SUTTER.COM

Where Science Gets Social.

AAAS.ORG/COMMUNITY

AAAS' Member Community is a one-stop destination for scientists and STEM enthusiasts alike. It's "Where Science Gets Social": a community where facts matter, ideas are big and there's always a reason to come hang out, share, discuss and explore.

Member COMMUNITY
AAAS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

OCEANOGRAPHY

Coral reef islands adjust to rising sea

Gravel-covered islands that develop naturally atop coral reefs have long provided opportunities for human habitation in the tropics, but there is widespread concern that such communities may be at risk because of sea level rise. However, Masselink *et al*. suggest that these islands may be able to adjust dynamically to changes in sea level. Numerical modeling of sediment transport on these islands during limited to moderate sea level rise indicated that foundering beneath the waves is not inevitable. Instead, sediment redistribution may simply change island geography and topography. Although the authors note that this process would be societally problematic, it may at least preserve the short-term habitability of coral reef islands. —KVH *Sci. Adv.* 10.1126/sciadv.aay3656 (2020).

Funafuti atoll, the capital of the Pacific island nation Tuvalu, is threatened by rising sea level, which may also change patterns of sediment deposition.

PHOTO: FIONA GOODALL/GETTY IMAGES FOR LUMIX

MESOSCOPIC PHYSICS

Coherent electronic waves

The wave nature of conducting electrons in solids can be revealed through interference effects. In layered materials, these effects are most often seen in in-plane transport. By contrast, Putzke *et al*. studied electronic transport perpendicular to the conductive layers in the ultraclean delafossites $PdCoO_2$ and $PtCoO_2$. When an in-plane magnetic field was applied, the electrical resistance exhibited periodic oscillations as a function of field magnitude. The findings can be explained through a model that requires that the electronic waves remain coherent over macroscopic distances. —JS

Science, this issue p. 1234

CELL CYCLE

How cells monitor mitogen availability

Classical experiments indicated that cells sense the mitogens or growth factors that control cell division within a limited window during the cell cycle. Min *et al*. reexamined this issue with high-throughput live-cell imaging and temporally controlled perturbations to more closely monitor dynamic signal processing. Human epithelial cells in culture integrated the mitogenic signals sensed throughout the cell cycle. One important factor was the control of translation rates, which influenced the amount of cyclin D, thus regulating proliferation. The results may also help to explain how cells maintain a uniform size. —LBR

Science, this issue p. 1261

PROTEIN FOLDING

Shedding light on disordered proteins

Disordered proteins often fold as they bind to a partner protein. There could be many different molecular trajectories between the unbound proteins and the bound complex. Most methods to measure transition paths rely on monitoring a single distance, making it difficult to resolve complex pathways. Kim and Chung used fast three-color single-molecule Foster resonance energy transfer (FRET) to simultaneously probe distance changes between the two ends of an unfolded protein and between each end and a probe on the partner protein. They show that binding can be initiated by diverse conformations and that the molecules are held together by non-native interactions as the disordered protein folds. This allows the association to be diffusion limited because most collisions lead to binding. —VV

Science, this issue p. 1253

NANOPHOTONICS

Light under compression

The ability to confine light to volumes much smaller than the wavelength produces high electromagnetic fields that can then be exploited in chemical and biological sensing and detection applications. Using silver nanocubes placed on a graphene surface, Epstein *et al*. developed a single, nanometer-scale acoustic graphene plasmon cavity device that can confine mid-infrared and terahertz radiation with mode volume confinement factors of 5×10^{10}. With the response

being dependent on the size of the nanocube and electrically tunable, the results demonstrate a powerful platform with which to develop sensors in what has been a challenging wavelength regime where molecular fingerprints reside. —ISO

Science, this issue p. 1219

MOLECULAR MACHINES

Ten rings on one axle

Rotaxanes consist of molecular rings threaded on a central axle. Most approaches to their synthesis have focused on introducing a single ring per axle. Qiu *et al*. now report a systematic approach to threading up to 10 adjacent rings consecutively. The axle's end groups were constructed to attract free-floating rings when reduced and then to push those rings toward the center upon oxidation. Products of each successive reduction-oxidation cycle were characterized by nuclear magnetic resonance spectroscopy and mass spectrometry. —JSY

Science, this issue p. 1247

INSECT LOCOMOTION

Robot on a wire

Studying the free flight of insects using artificial systems such as tethered flight or flight in confined environments can cause the insect to fly in an unnatural way. Pannequin *et al.* report a cable robot consisting of an open cage mounted with cameras, which they call "lab-on-cables." The cable robot moves automatically with the flying insect by using reactive controllers to track the insect's flight in natural conditions with limited disturbance. Using this system, the researchers were able to study *Agrotis ipsilon* moths flying freely at speeds of up to 3 meters per second. —MML

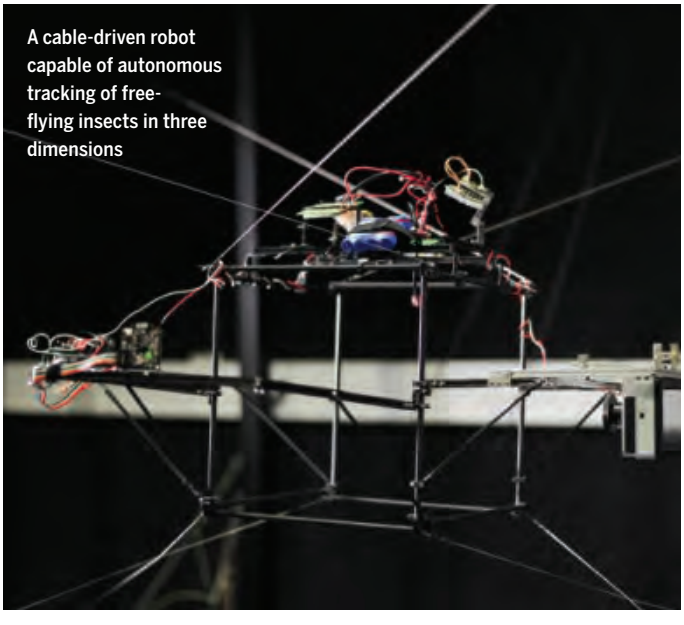
Sci. Robot. **5**, eabb2890 (2020).

ICE SHEETS

Taking stock of our losses

Earth's ice sheets are melting and sea levels are rising, so it behooves us to understand better which climate processes are responsible for how much of the mass loss. Smith *et al*. estimated grounded and floating ice mass change for the Greenland and Antarctic ice sheets from 2003 to 2019 using satellite laser altimetry data from NASA's ICESat and ICESat-2 satellites. They show how changing ice flow, melting, and precipitation affect different regions of ice and estimate that grounded-ice loss averaged close to 320 gigatons per year over that period and contributed 14 millimeters to sea level rise. —HJS

Science, this issue p. 1239



A cable-driven robot capable of autonomous tracking of free-flying insects in three dimensions

IN OTHER JOURNALS

Edited by **Caroline Ash** *and* **Jesse Smith**

GENETICS

Big cat genomics

The lion may be the king of beasts, but over the past ~14,000 years, their range has decreased substantially and many current populations are in decline. To understand the historical population genomics of lions, de Manuel *et al.* sequenced cave lion remains from ~30,000 years ago, several historic specimens from extinct populations over the past ~500 years, and samples from extant populations. From this, they were able to reconstruct relationships among lion populations over time. No evidence of gene flow between cave lions and recent lion populations was detected. Modern lion groups that diverged into northern and southern lineages show evidence of admixture, especially in extant central African populations. These data could be valuable not only because they indicate the relationships between modern and extinct populations but also because they reveal a historical level of inbreeding that could be relevant to conservation efforts. —LMZ

Proc. Natl. Acad. Sci. U.S.A. **117**, 10927 (2020).

SIGNALING

Controlling blood flow in the liver

High blood pressure in the liver created by fibrosis causes serious clinical illness. Blood pressure is regulated by nitric oxide, but this is depleted if hepatic sinusoidal endothelial cells are damaged. The scaffold protein β-arrestin 2 (β-Arr2) regulates the activity of endothelial cell nitric oxide synthase (eNOS) in rodent liver. Liu *et al.* found that in rodents, β-Arr2, which serves several functions in G protein–coupled receptor signaling, helps to control eNOS at endothelin receptors. β-Arr2 promoted the association of another protein called GIT1 with eNOS, which activates eNOS. Apparently, β-Arr2 and other signaling components form a signalsome near the endothelin receptors, which regulates eNOS function. Disruption of this signaling contributes to the liver injury and portal hypertension that cause lethal liver disease in humans. —LBR

Proc. Natl. Acad. Sci. U.S.A. **117**, 11483 (2020).

MIGRATION

Should I stay or should I go?

The global upheaval of seasonal bird migration is an astonishing phenomenon. The European blackcap is a species of small songbird that shows great variation in wanderlust. Delmore *et al.* collected DNA-sequencing data from blackcaps across their breeding range to determine the genetic basis for variations in migratory behavior. The impulse to migrate and the distance and direction of travel are encoded in small genomic regions. Polymorphisms were found in genes for transcription factors such as Clock, Npas2, and Bma1, which are associated with circadian rhythm. A few genes linked to traits as varied

PHOTOS (FROM LEFT): DOMINIQUE MARTINEZ; SWAPAN BANIK / ALAMY STOCK PHOTO

Gir lions are a remnant population of a few hundred northern lions, mostly found in Africa, that are just surviving in India.

as hyperphagia and fat deposition, learning and memory, and wing length also appear to be under strong selection to rapidly evolve new behaviors. The suite of blackcap migratory genes is distinctive, and it appears that there are many ways to genetically program migration in birds. —CA

eLife **9**, e54462 (2020).

NANOMATERIALS

Pattering nanoparticles with DNA origami

Complementary DNA strands can be used to assemble nanoparticles through specific connections, but creating multivalent directional connections is still challenging. Xiong *et al.* used DNA origami—a two-dimensional open square and a three-dimensional tetrahedron framework—to position DNA linkers on nanoparticles. These molecular stamping, or MOST, frames were "inked" with single-stranded DNA that transferred onto a gold nanoparticle bound inside the frame. On release, the particles could then undergo complementary strand binding with smaller gold nanoparticles to form clusters. By using different inks within the frame, gold particles of different sizes could be assembled onto the central particle. The angles between the particles in these heterogeneous clusters could also be controlled through steric effects. —PDS

ACS Nano. 10.1021/acsnano.0c00607 (2020).

PSYCHOLOGY

Procedural justice improves policing

Research on procedural justice emphasizes the importance of treating people fairly regardless of the outcome they receive. Procedural justice strategies include increased transparency, communication, and responsiveness to civilian concerns. Wood *et al.* examined whether the introduction of a procedural justice training program for police officers in Chicago had an impact on encounters between police and civilians. It was found that fewer complaints were registered against officers who received the training, and trained officers were less likely to use force during civilian encounters. These findings have implications for designing scalable interventions that build police legitimacy. —TSR

Proc. Natl. Acad. Sci. U.S.A. **117**, 9815 (2020).

MACHINE LEARNING

Learning macro from micro

Determining atomic structural correlations in condensed-phase systems is crucial for understanding material properties and their behavior at the macroscale. It represents one of the central challenges in modern statistical mechanics because of the complex collective behavior emerging from microscopic many-body interactions. Using two classical condensed-phase models, a Lennard-Jones system and a hard-sphere fluid, Craven *et al.* show that machine learning methods trained on a set of optimally short molecular dynamics simulations can predict radial distribution functions with increased accuracy by an order of magnitude or even greater compared with traditional analytical approaches. The proposed methodology is general and could be applied more broadly across diverse condensed-phase systems. —YS

J. Phys. Chem. Lett. 10.1021/acs.jpclett.0c00627 (2020).

NEUROSCIENCE

Inositol-triphosphate receptors in axons

Inositol-triphosphate (IP_3) receptors regulate the intracellular calcium concentration in the somatodendritic compartment of central neurons. Whether axons also possess functional IP_3 receptors and what impact their activation might have are not known. Cerebellar Purkinje cells offer an ideal model because they contain a high level of IP_3 receptors. Using chromophore tags that release IP_3 when irradiated, Gomez *et al.* found that functional IP_3 receptors are present in the entire axon. Different axon regions displayed different IP_3-producing pathways, and IP_3 receptor activation had different consequences depending on receptor localization. For instance, IP_3 receptor activation in synaptic terminals caused neurotransmitter release, and receptor activation in the axon initial segment blocked action potential firing. IP_3 receptor–linked signaling pathways may therefore be important in controlling axon functions. —PRS

Proc. Natl. Acad. Sci. U.S.A. **117**, 11097 (2020).

REVIEW SUMMARY

NEUROSCIENCE

Linking glutamate receptor movements and synapse function

Laurent Groc* and Daniel Choquet*

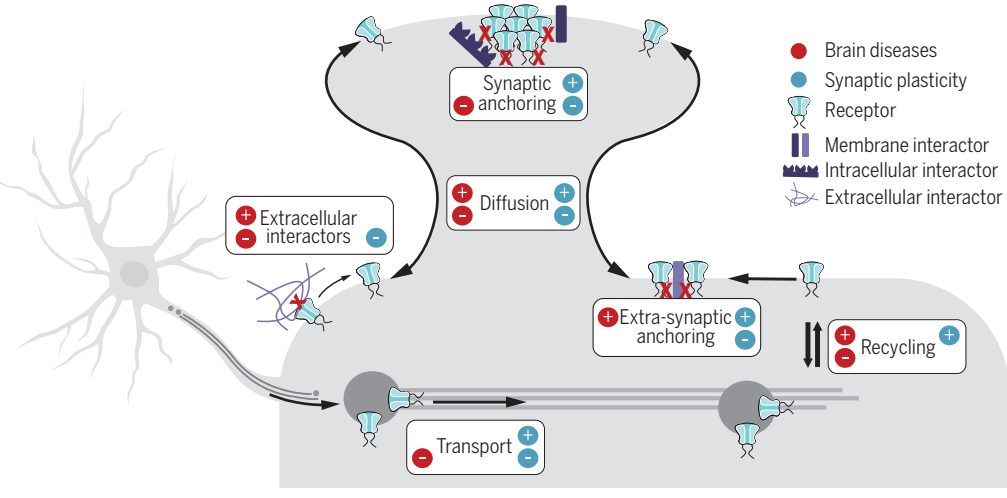
BACKGROUND: Since it was established that the cognitive brain is formed mostly by an interconnected network of neurons that communicate at contact sites termed synapses, intense research has aimed at identifying their molecular composition and physiological roles. The discovery that the efficacy of synaptic transmission can be modified by neuronal activity has undoubtedly been a major step in understanding brain function. The various forms of activity-dependent synaptic plasticity were early on proposed to play central roles in brain adaptation, learning, and memory. This motivated neurophysiologists to understand the mechanisms of synaptic plasticity, initially within the sole framework of the quantal properties of transmitter release, largely ignoring the cell biology revolution that was occurring in parallel. In the 1970s, at the same time that synaptic plasticity was discovered, the fluidity of cell membranes was established. Surprisingly, these contemporary findings seldom crossed paths. As cell biologists established the major roles of receptor trafficking in cell function, neurophysiologists still largely viewed synapse function as based on unitary receptor properties and control of transmitter release. It has been only about 20 years since the two fields cross-fertilized and the regulation of receptor movements into and out of synapses emerged as a fundamental mechanism for synaptic plasticity.

ADVANCES: Largely based on the development of imaging approaches, including single-molecule tracking, receptors have been demonstrated to undergo a variety of movements, from long-range rapid motor-based intracellular transport, to short-range Brownian surface diffusion, and intercompartment exchange by membrane trafficking. For efficient synaptic transmission, receptors must accumulate in front of neurotransmitter release sites. This is accomplished through a set of interactions with intracellular scaffold proteins, transmembrane auxiliary subunits, or adhesion proteins and other extracellular elements. This duality of receptor movements and stabilization has led to the important concept that the number of functionally responsive receptors at synapses results from the interplay between reversible receptor stabilization and dynamic equilibrium between pools of receptors in the synaptic, extrasynaptic, and intracellular compartments. Coarse receptor distribution along dendrites is largely achieved by intracellular transport. Because exchange of receptors between surface and intracellular compartments seems to occur largely at extrasynaptic sites, reversible surface receptor diffusion trapping at synapses has emerged as a central mechanism to control their availability for synaptic activation. Receptor stabilization and movements are all profoundly regulated by short- and long-term neuronal activity patterns. Reciprocally, evidence has accumulated that receptor movements participate in many forms of synaptic plasticity. Notably, altered receptor movements are observed in many neurodevelopmental, psychiatric, or neurodegenerative pathological models as indicated in the figure [the + and – signs indicate the reported positive and negative modulation of the indicated trafficking and stabilization processes during either normal (blue) or pathological (red) synaptic function]. Whether altered receptor trafficking represents the primum movens of some neurological diseases remains to be established, but is certainly an attractive hypothesis.

ON OUR WEBSITE

Read the full article at https://dx.doi.org/10.1126/science.aay4631

OUTLOOK: Most receptor trafficking studies have been performed in reduced experimental systems such as neuronal cultures. This has limited our understanding of the physiological impact of these processes. The development of brighter and smaller probes together with new imaging modalities are on the verge of allowing routine measurement of receptor movements in more physiological settings such as brain slices and in vivo. There is little doubt that qualitatively comparable trafficking modalities will be identified. Reciprocally, tools are being developed to control the various types of receptor movements, from blocking surface diffusion by receptor cross-linking to stopping receptor exocytosis with light-activated toxins. Often, these trafficking tools do not impair basic synaptic function, because resilience of the synapse to trafficking alterations is high owing to the amount of available receptors, as well as the trapping capacities and nanoscale organization of the synapse. Combining measurement and control of receptor movements will not only allow better understanding of their contribution to synaptic and neuronal function but also provide valuable tools for identifying the role of synaptic plasticity in higher brain functions. Controlling receptor movements or stabilization may eventually represent an alternative therapeutic strategy to receptor activity modulation approaches in a variety of synaptic and network-based brain diseases. ■



Neurotransmitter receptors undergo a variety of large- and small-scale movements. Movements of large amplitude constantly reshuffle the receptor distribution (e.g., surface diffusion and intracellular transport). Movements at interfaces (e.g., between synaptic and extrasynaptic sites, between intracellular and surface compartments) are of small amplitude but have huge functional impacts. Each of these movements is highly regulated and finely tuned in physiological and pathological conditions.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: laurent.groc@u-bordeaux.fr (L.G.); daniel.choquet@u-bordeaux.fr (D.C.)
Cite this article as L. Groc, D. Choquet, *Science* 368, eaay4631 (2020). DOI: 10.1126/science.aay4631

RESEARCH ARTICLES

CELL BIOLOGY

Membrane-proximal F-actin restricts local membrane protrusions and directs cell migration

Anjali Bisaria[1]*, Arnold Hayer[1]†, Damien Garbett[1], Daniel Cohen[2]‡, Tobias Meyer[1,3]*

Cell migration is driven by local membrane protrusion through directed polymerization of F-actin at the front. However, F-actin next to the plasma membrane also tethers the membrane and thus resists outgoing protrusions. Here, we developed a fluorescent reporter to monitor changes in the density of membrane-proximal F-actin (MPA) during membrane protrusion and cell migration. Unlike the total F-actin concentration, which was high in the front of migrating cells, MPA density was low in the front and high in the back. Back-to-front MPA density gradients were controlled by higher cofilin-mediated turnover of F-actin in the front. Furthermore, nascent membrane protrusions selectively extended outward from areas where MPA density was reduced. Thus, locally low MPA density directs local membrane protrusions and stabilizes cell polarization during cell migration.

The cell cortex of eukaryotic cells is composed of a network of F-actin filaments below the plasma membrane (PM) (*1*). Changes in the organization, thickness, and contractility of the cortex modulate cell shape and are important for a variety of cell processes such as cell signaling and cell migration (*2*–*4*). The PM is reversibly attached to the cortex by several linker mechanisms (fig. S1A), including the binding of ezrin, radixin, and moesin (ERM) proteins to both membrane and F-actin (*5*). Membrane attachment through linker proteins is restricted to the outermost layer of the cell cortex, here called membrane-proximal F-actin (MPA), with the strength of attachment believed to be regulated by the distance or density of MPA (*6*).

The actin cortex has a role in directing migrating cells in part through regulating membrane tension (*7*, *8*). The activities controlling severing, nucleation, and contraction of F-actin are polarized and lead to both higher F-actin turnover and high F-actin density in the front (*9*, *10*), whereas ERM proteins are enriched in the back (*11*). A priori, one would expect that the front of a cell should have high MPA density because of the higher F-actin concentration in lamellipodia (*12*). If true, however, increased MPA density could lead to stronger membrane attachment in the front, which would be expected to oppose rather than promote protrusion. It is unclear whether MPA density is polarized, how local MPA density is regulated, and if MPA density regulates local protrusion in the context of cell migration.

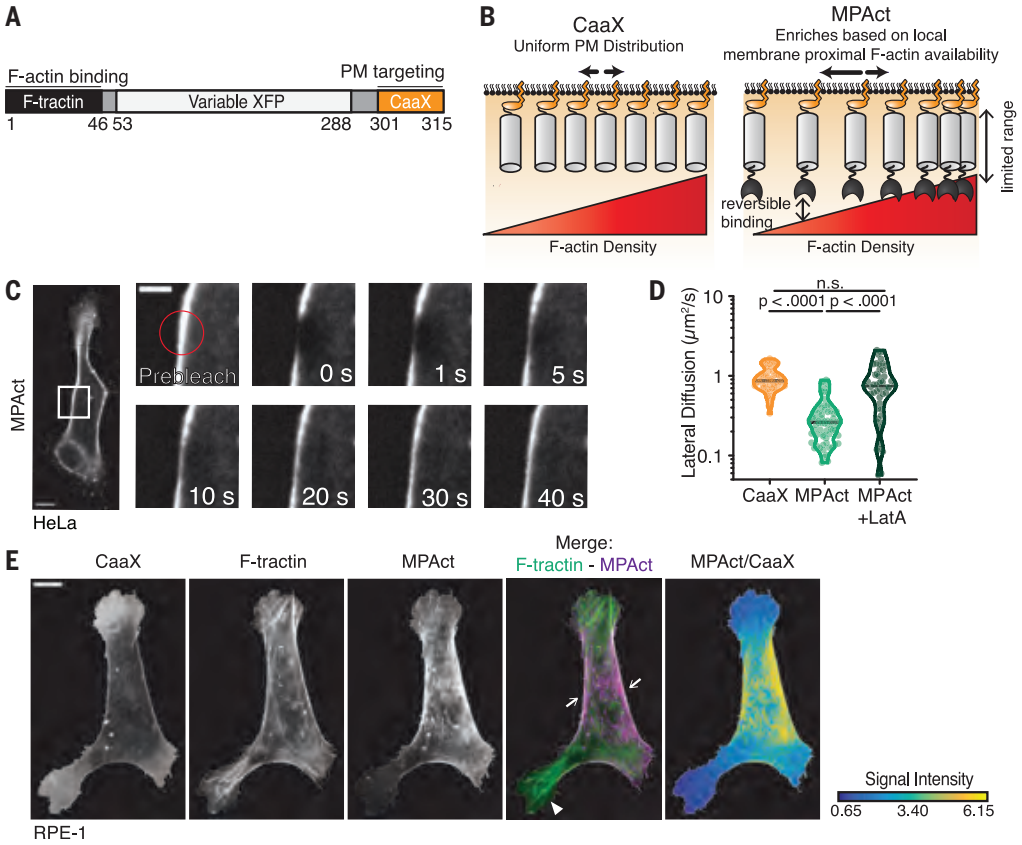


Fig. 1. Development of a fluorescent reporter that measures the local density of membrane-proximal F-actin. (**A**) Design of the MPAct reporter. (**B**) Ratiometric analysis of MPA density. The MPAct signal was normalized by the CaaX PM marker. (**C**) FRAP analysis of lateral membranes of RPE-1 cells expressing MPAct-mCitrine. Scale bar, 5 μm. (**D**) Diffusion coefficients calculated from FRAP experiments in RPE-1 with and without 1 μM LatA. *N* = 52, 69, and 54, respectively, from two independent experiments. Medians are shown. Dotted lines are quartiles. *P* values were derived from unpaired *t* test. (**E**) RPE-1 cells expressing MPAct-mCitrine (magenta), CFP-CaaX, and F-tractin-mCherry (green). For all figures, scale bars are 10 μm unless marked and ratiometric images are shown in Parula color scale from blue (low) to yellow (high).

[1]Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA, USA. [2]Department of Biology, Stanford University, Stanford, CA, USA. [3]Department of Cell and Developmental Biology, Weill Cornell Medical College, New York, NY 10065, USA.
***Corresponding author. Email: abisaria@stanford.edu (A.B.); tom4003@med.cornell.edu (T.M.)** †Present address: Department of Biology, McGill University, Montréal, Québec, Canada. ‡Present address: Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, USA.

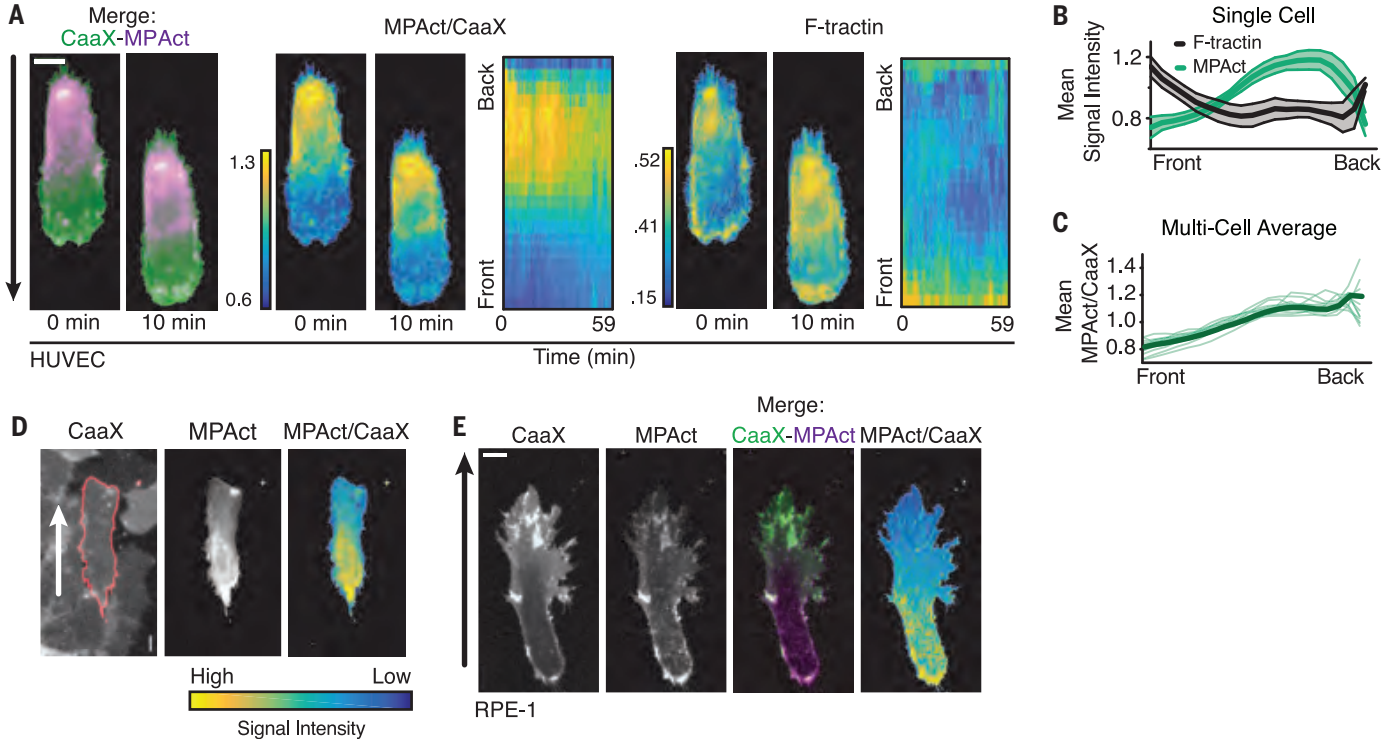


Fig. 2. Migrating cells have stable back-to-front MPA density gradients, with MPA density being lowest in the front independent of cell type and migration mode. (**A**) HT-HUVECs expressing MPAct-mCitrine (magenta) and iRFP-CaaX (green) migrating on 20-μm fibronectin stripes (top). MPAct/CaaX ratio (middle) and F-actin marker F-tractin (right) images are shown at the same time points. Kymographs of activity changes compare front-to-back location. Time points are 2 min. (**B**) Quantification of activity profile changes over 1 hour for the cell in (A). Error bars indicate SD. (**C**) Average gradients of the MPAct/CaaX ratio for 14 individual HT-HUVECs (green) and group average (black) migrating on stripes. (**D**) Example of MPAct/CaaX ratio image in collectively migrating HT-HUVEC monolayers with mosaic MPAct expression. (**E**) Image section of RPE-1 cells stably expressing iRFP-CaaX (green) and MPAct-mCitrine (magenta) migrating in a 0.5 mg/ml collagen matrix. Scale of MPAct/CaaX is normalized for maximal signal range. Migration directions are indicated by arrows.

Cells migrating in one, two, and three dimensions have stable MPA gradients with low MPA density in the front

To test directly for spatial heterogeneity of MPA density, we developed a membrane-proximal F-actin reporter (MPAct) to monitor the local density of MPA. We designed the MPAct reporter by linking a fast-diffusing PM anchor (CaaX) (*13*) to a low-affinity F-actin–binding domain (F-tractin) (*14*) and a fluorescent protein (XFP) (Fig. 1A). The membrane anchor allowed the reporter to diffuse rapidly along the inner leaflet of the PM and to enrich in regions with higher densities of F-actin near the membrane (Fig. 1, A and B). Rapidly reversible binding of the reporter to F-actin ensured that fast changes in local F-actin density could be measured, because the cell cortex is continuously turned over (*15*). The reporter remains inserted into the PM with a limited search radius that restricts F-actin binding to a membrane-proximal attachment zone within 10 nm from the membrane, much shorter than the ~200-nm-thick actin cortex (*16*).

We validated that the reporter was exclusively membrane localized and measured local F-actin density near the surface membrane (figs. S1 and S2). To determine the relative local MPA density along the PM surface, a ratiometric MPAct signal was measured by dividing the MPAct intensity over the intensity of a CaaX PM marker (here called MPAct/CaaX or MPAct signal) (Figs. 1E and 2, A and B).

In migrating hTERT-immortalized human umbilical vein endothelial cells (HT-HUVECs), we unexpectedly observed an MPAct signal gradient that was low in the front of migrating cells (Fig. 2, A to C, and movie S1). This contrasted with an overall high global F-actin density in the same front region (Fig. 2A, right) (*12*, *17*). Kymograph analysis of HT-HUVEC and retinal pigment epithelium (RPE-1) cells showed that MPAct signal gradients were stable during migration (Fig. 2A and fig. S3, A to E). Furthermore, the same MPAct signal gradients were observed in single-migrating and collectively migrating cells, in cells migrating on collagen or fibronectin substrates, in cells in different migration contexts (one, two, and three dimensions), and in cells visualized after using different microscope modalities (confocal, total internal reflection fluorescence, epifluorescence) (Fig. 2, D and E, and figs. S2 and S3, F to J). Lower F-actin binding in the front resulted in a slower fluorescence recovery after photobleaching (FRAP) recovery in the back versus front (fig. S4). We observed the same back-to-front MPAct gradient when we varied the type of F-actin–binding motif or membrane anchor in the MPAct reporter as when we used a different PM marker as a reference for normalization (fig. S5). Finally, the MPAct gradient was identical when we varied the length and flexibility of the linker in the MPAct reporter (fig. S6), suggesting that the MPAct reporter measures the averaged local MPA density. Thus, migrating cells have an underlying gradient of MPA density, with MPA density lowest in the front.

MPA density is locally decreased before membrane protrusion during de novo cell polarization

When we monitored RPE-1 cells migrating on collagen-coated glass slides, cells frequently repolarized and the MPAct signal gradient

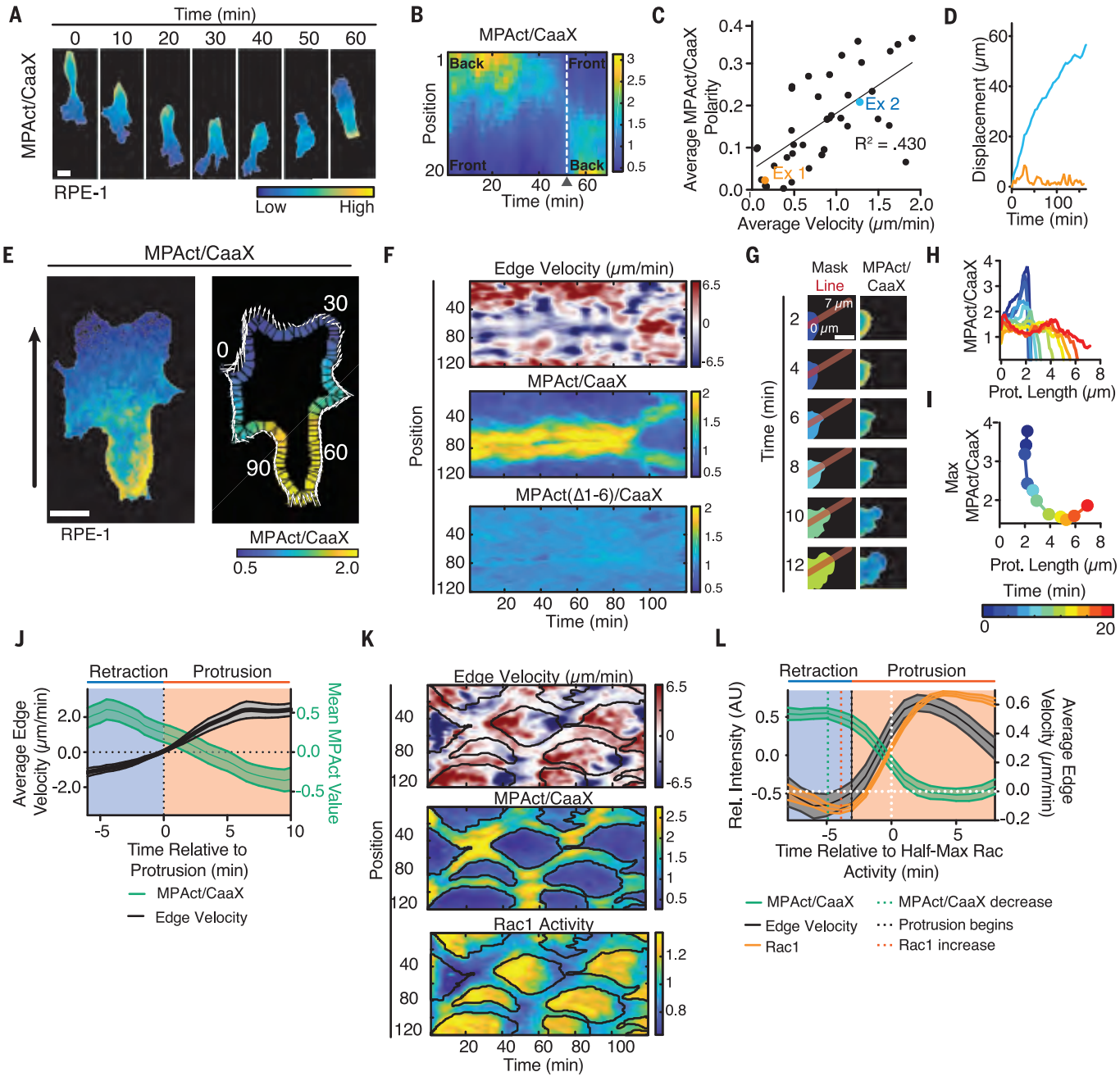


Fig. 3. MPA density is locally lowered before Rac activation and membrane protrusion. (**A** and **B**) Repolarization of migrating RPE-1, MPAct/CaaX image series (A) and kymograph analysis (B). (**C** and **D**) Mean MPAct/CaaX gradient steepness versus mean velocity for HT-HUVECs migrating on stripes over periods >1 hour. Gradient steepness was calculated as the ratio of MPAct signal between back and front 10% of cell. Average velocity is total displacement/time. Line of best fit is shown. *N* = 43 cells from two independent experiments. (D) Centroid displacement for the two cells marked in (C). (**E**) Edge parametrization. Local protrusion direction and distance are shown as white arrows and global direction as a black arrow. Edge correlation analysis from (*18*). (**F**) Comparison of local edge velocity (top), MPAct/CaaX (middle), and MPAct(Δ1-6)/CaaX (bottom) signals versus time. Images were taken every 1 min. (**G** to **I**) MPAct signal change during local retraction and protrusion. Red line shows profile direction (G). Shown are the analysis of MPAct/CaaX line profile over time (H) and maximal MPAct signal as a function of time and protrusion length (I). Scale bar, 5 μm. (**J**) Comparison of mean protrusion edge velocity and local MPAct signal change. Traces are aligned to time of protrusion. Shown are means and 95% confidence intervals (CIs) from *N* = 25 repolarization events. (**K**) Kymograph of normalized edge velocity (top), MPAct/membrane (middle), and Raichu-Rac1 FRET (bottom) in RPE-1. Edge analysis was as in (E) and (F). Black outlines mark areas of low MPAct signals. (**L**) In silico activity alignment to half-maximal Rac activation comparing edge velocity (black, right *y*-axis), MPAct (green, left *y*-axis), and Raichu-Rac1 (orange, left *y*-axis) signal time courses. Data are shown as means and 95% CI from *N* = 30 events and three independent experiments.

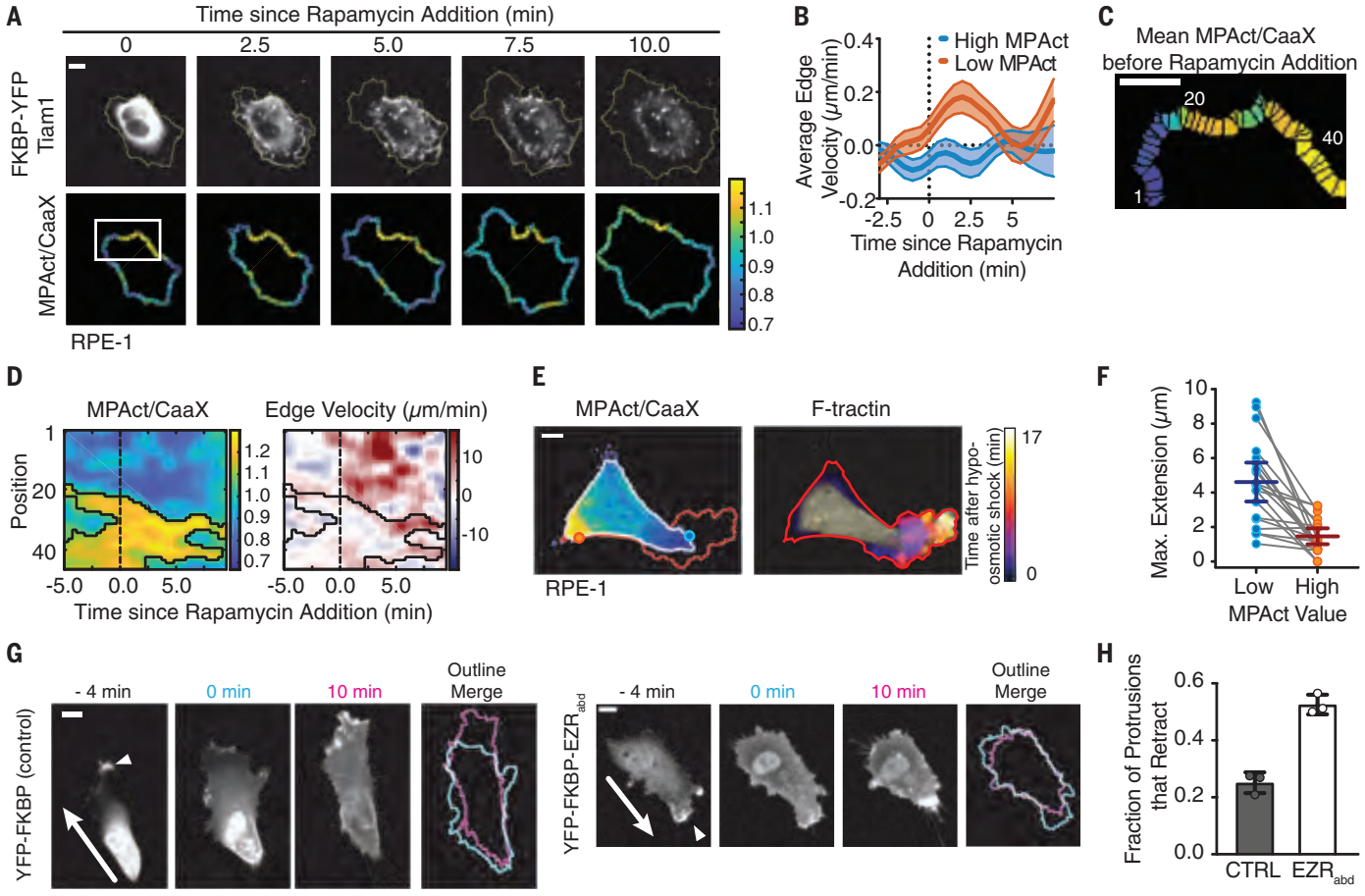


Fig. 4. Local sites with low MPA density direct local membrane protrusions by Rac-mediated actin polymerization and osmotic pressure. (**A** to **D**) Cell response after uniform acute Rac-GEF activation (TIAM). Membrane outline changes are shown as a yellow periphery (A). Induced edge velocity comparing areas of high (red) and low (blue) MPAct signal. N = 27 cells from two independent experiments (B). Kymograph focuses on the white box (C) from the image in (A) comparing MPAct/CaaX signal (left) and edge velocity (right). Black outlines mark areas of high MPAct signals. Scale bar, 5 μm. (**E**) (Left) MPAct/CaaX image before hypo-osmotic shock. Maximum future expansion is shown in red and initial cell area in white. (Right) F-tractin image showing progressive expansion after hypo-osmotic shock. (**F**) Hypo-osmotic membrane extension over 10 min for high versus low MPAct areas. Data from the same cell are connected by gray lines. Mean and 95% CI shown from N = 20 cells. (**G**) Control FKBP-YFP (left) and FKBP-mCitrine-Ezrin$_{abd}$ (right) during rapamycin addition. Active protrusion is marked by a white arrow, and outlines at 0 and 10 min are shown. (**H**) Mean fraction of retracting protrusions from N = 109 (CTRL) and N = 121 (EZR$_{abd}$) cells from three separate experiments. Mean and SD are shown.

formed in the new direction of migration (Fig. 3, A and B, and movie S2). By contrast, a mutated MPAct reporter that could not bind F-actin failed to form gradients (fig. S4, D to F, and movie S2). Furthermore, the steepness of the MPAct gradient and average migration velocity were correlated (R^2 = 0.430, N = 44 cells from two independent experiments; Fig. 3, C and D, and fig. S7, A and B). Thus, cells may need to decrease MPA density locally to protrude membranes, which in turn may direct and promote cell migration.

To test for a suppressive role of local MPA density in membrane protrusion, we parameterized the edge of cells and correlated local MPAct signals with protrusion and/or retraction rates along the cell periphery (Fig. 3E). This analysis showed a close spatial anticorrelation between low MPAct level and protrusion that was not recapitulated by the control MPAct(Δ1-6) reporter construct (Fig. 3F and fig. S7, C and D). An in silico alignment of the local edge velocity versus the local MPAct signal change during repolarization showed a decrease of MPAct signals ~4 min before protrusion (Fig. 3, G to J, and fig. S7, E to G). These results suggested that locally high MPA density inhibits membrane protrusions and that MPA density must be lowered to promote membrane protrusion.

Local MPA density controls Rac-induced local membrane protrusion

Rac activity increases F-actin polymerization and triggers broad membrane protrusions at the front of many migrating cell types (*18*). Using a fluorescence resonance energy transfer (FRET) reporter for Rac activity (*19*), we found that areas of high Rac activity and protrusion-retraction cycles were closely confined to areas of low MPAct signals (Fig. 3K and fig. S7, H to J). When aligning MPAct signals in silico to the time of half-maximal Rac activation, Rac activity increased only after a reduction of MPAct signals (Fig. 3L). The prior reduction in MPA density suggested that high MPA density inhibits both local protrusion and Rac activation.

To directly test the hypothesis that high MPA density inhibits Rac-driven protrusion, we used a Lyn-FRB and Tiam1-FKBP activation system to activate endogenous Rac by rapamycin-induced translocation of the

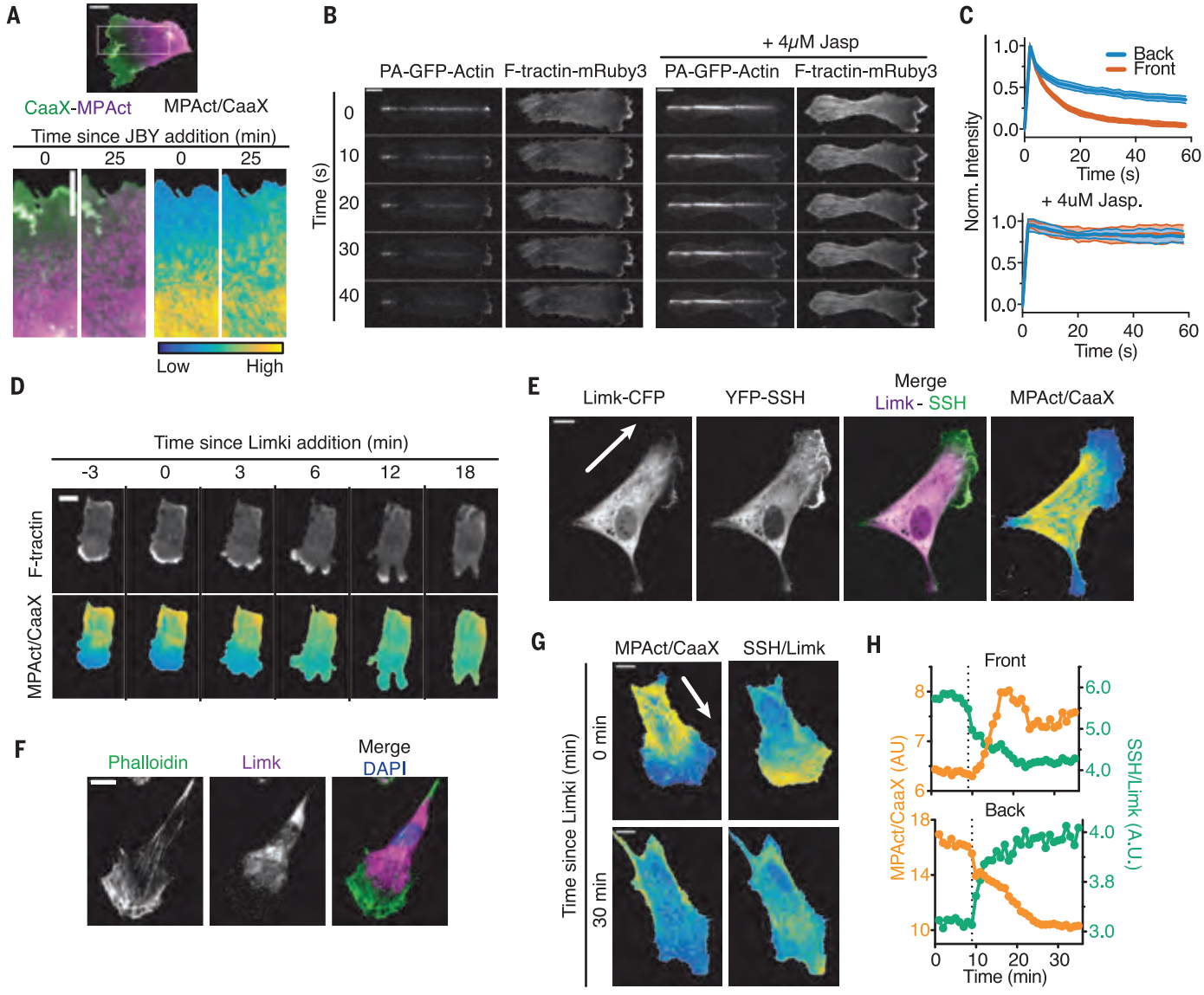


Fig. 5. Higher cofilin-mediated F-actin disassembly in the front versus back regulates the MPA density gradient during cell migration. (**A**) (Top) Migrating RPE-1 before treatment with the JBY cocktail. Boxed region is shown below. (Bottom) CaaX (green), MPAct (magenta), and MPAct/CaaX images shown before and after the addition of JBY cocktail. (**B**) PA-GFP-actin and F-tractin dynamics after a PA-GFP-Actin line photoconversion in migrating RPE-1, in control (left), and after the addition of 4 μM jasplakinolide to stabilize F-actin (right). (**C**) Mean PA-GFP-actin fluorescence loss at the front (red) and back (blue) of migrating RPE-1 WT (top, *N* = 44 cells) or after the addition of 4 μM jasplakinolide (bottom, *N* = 9 cells). Means and 95% CIs are shown. (**D**) F-tractin (top) and MPAct/CaaX (bottom) signal changes after the addition of 20 μM LIMKi 3. (**E**) RPE-1 expressing Limk-CFP (magenta), YFP-SSH (green), MPact-mRuby3, and iRFP-CaaX. (**F**) RPE-1 stained with phalloidin, anti-αLimk antibody, and Hoechst. (**G**) Distribution of MPAct/CaaX versus SSH/Limk before and after the addition of 20 μM LIMKi (typical example, *N* = 10 cells). (**H**) MPAct/CaaX (orange, left *y*-axis) and SSH/Limk (green, right *y*-axis) signal changes in front and back for cell in (G). Dashed line indicates LIMKi addition.

Rac GEF Tiam1 to the PM (*20*) (fig. S8A). Rapid and uniform Tiam1 activation using this system triggered selective protrusions only from areas with low MPAct signals (Fig. 4, A to D). Protrusions can also be triggered when cells are subjected to hypo-osmotic shock (*21*). We lowered osmolarity by water addition and observed rapid membrane protrusions selectively from areas of low MPAct signals (Fig. 4, E and F, and fig. S8B). Thus, multiple types of protrusions are directed to low MPA areas.

To directly test whether high MPA density suppresses local protrusion, we engineered a system to acutely increase MPA density using a FKBP construct containing the high-affinity [~40 nM (*22*)] F-actin–binding domain of Ezrin ($Ezrin_{abd}$; fig. S8C). Upon addition of rapamycin, FKBP-$Ezrin_{abd}$ was recruited to the membrane and increased MPAct signals, as expected (fig. S8, D to F). Compared with a control construct (Fig. 4G, left), recruitment of FKBP-$Erzin_{abd}$ (Fig. 4G, right) triggered increased membrane retraction (Fig. 4H). Thus, high MPA density

prohibits membrane protrusions and lowering MPA density is required to sensitize areas for membrane protrusion.

High rates of cofilin-regulated MPA disassembly control locally low MPA density in the front

We next investigated how cells reduce MPA density in the front despite a higher overall F-actin density there. Low MPA density could be maintained dynamically by relatively higher F-actin disassembly, lower actin polymerization, lower numbers of F-actin linker proteins, or lower actomyosin contraction rates in the MPA attachment zone in the front (fig. S9A). We performed a series of rapid perturbation experiments using inhibitors of F-actin regulators to distinguish potential mechanisms and found that most inhibitors showed minimal relative short-term effects (fig. S9, B to E). The low MPAct signal in the front was lost when we treated cells with a jasplakinolide, blebbistatin, and Y-27632 (JBY) cocktail to inhibit F-actin disassembly and contraction without preventing polymerization (Fig. 5A and fig. S9, F and G). In an orthogonal experiment, we used photoactivation of a green fluorescent protein (GFP)–tagged actin and monitored loss of the fluorescent signal along the axis of migration. Consistent with a relatively faster F-actin disassembly in the front (*23*), the GFP-actin fluorescence was lost faster in the front compared with the back (Fig. 5, B and C, and fig. S9, H to J). Thus, a primary mechanism controlling locally low MPA density is continued high F-actin disassembly in the front.

The ubiquitous F-actin–severing protein cofilin is known to have higher activity in the front of migrating cells (*24*). We hypothesized that cofilin activity could have a distinct effect on MPA versus global F-actin at the front. Cofilin is activated by slingshot phosphatase (SSH)–mediated dephosphorylation and inactivated by Limk-mediated phosphorylation (fig. S9L) (*24*). Treatment with the LIMKi 3 inhibitor triggered a rapid equilibration of the MPAct gradient between the front and back (Fig. 5D, fig. S9K, and movie S3) and also caused an expected transient increase in the size of membrane protrusions by making more G-actin available for actin polymerization (fig. S9M) (*24*). Among several inhibitors of F-actin–regulatory processes that we tested (fig. S9, B to G, and K), only inhibition of Limk and F-actin turnover resulted in a rapid loss of the MPAct gradient that preceded the overall cell polarity loss.

We confirmed that expressed fluorescently tagged SSH was enriched in the front, whereas expressed Limk was generally enriched in the back (Fig. 5, E to G, and fig. S9, N and O) (*9*). Upon inhibition of Limk, both the gradient of MPAct and the enrichment of SSH in the front of the cells were largely lost (Fig. 5H). Ratio images of SSH/Limk showed an inverse spatial gradient from front to back that closely paralleled the MPAct gradient during cell migration (Fig. 5, G and H). We conclude that cofilin-regulated local severing of MPA in the front is a critical mechanism establishing back-to-front gradients of MPA density.

Discussion

Here, we developed a ratiometric MPAct reporter that monitors relative differences in local MPA density. In all migrating cell types examined, the reporter showed an unexpected back-to-front gradient in MPA density, with the lowest MPA density at the front despite a higher actin polymerization rate and higher overall F-actin density there. Furthermore, under conditions in which cells locally protrude membranes, protrusions were invariably generated from sites with low MPA density. A consequence of having low MPA density and low MPA attachment is low membrane tension; therefore, local differences in MPA density can explain why membrane tension can be low in some areas of a cell and high in others (*25*). Thus, cells lower local MPA density to direct protrusions and cell movement, and stable back-to-front gradients of MPA density can function as a cytoskeletal "memory" in the face of fluctuating migratory signals to promote directional persistence.

Because cells have high global F-actin polymerization rates at the front, a correspondingly higher rate of MPA disassembly is required to sustain the lower MPA density in the same location. High cofilin activity in the front was required to establish and continuously maintain the MPA gradient during cell migration. It has previously been proposed that cofilin activation helps to initiate protrusion by providing free G-actin monomers (*24*). Our data suggest that cofilin has an additional role in lowering local MPA density and thereby direct protrusions.

Low MPA density can direct Rac and osmotically driven protrusions, suggesting that MPA gradients have an analogous role in cells that migrate by polymerizing actin forward or protruding membrane blebs (*26*). MPA gradients therefore specify "frontness" versus "backness" during migration regardless of the migration modality, and the steepness of MPA gradients can function as a persistence mechanism that regulates cell speed. Finally, the MPAct reporter will likely be useful for investigating a wide range of cellular processes, because MPA density is believed to regulate mechanotransduction (*27*) and the diffusion-reaction dynamics of G protein–coupled receptor and other membrane proteins (*28*).

REFERENCES AND NOTES

1. G. Salbreux, G. Charras, E. Paluch, *Trends Cell Biol.* **22**, 536–545 (2012).
2. K. J. Chalut, E. K. Paluch, *Dev. Cell* **38**, 571–573 (2016).
3. P. P. Ostrowski, S. Grinstein, S. A. Freeman, *Dev. Cell* **38**, 135–146 (2016).
4. S. A. Freeman *et al.*, *Cell* **172**, 305–317.e10 (2018).
5. M. Fritzsche, R. Thorogate, G. Charras, *Biophys. J.* **106**, 343–353 (2014).
6. M. P. Clausen, H. Colin-York, F. Schneider, C. Eggeling, M. Fritzsche, *J. Phys. D Appl. Phys.* **50**, 064002–064012 (2017).
7. Y. Liu *et al.*, *Blood* **119**, 445–453 (2012).
8. A. Diz-Muñoz *et al.*, *PLOS Biol.* **8**, e1000544 (2010).
9. M. Nishita *et al.*, *J. Cell Biol.* **171**, 349–359 (2005).
10. J. J. Bravo-Cordero, M. A. O. Magalhaes, R. J. Eddy, L. Hodgson, J. Condeelis, *Nat. Rev. Mol. Cell Biol.* **14**, 405–415 (2013).
11. Y.-J. Liu *et al.*, *Cell* **160**, 659–672 (2015).
12. A. Diz-Muñoz *et al.*, *BMC Biol.* **14**, 74 (2016).
13. W. D. Heo *et al.*, *Science* **314**, 1458–1461 (2006).
14. M. J. Schell, C. Erneux, R. F. Irvine, *J. Biol. Chem.* **276**, 37537–37546 (2001).
15. J.-Y. Tinevez *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 18581–18586 (2009).
16. A. G. Clark, K. Dierkes, E. K. Paluch, *Biophys. J.* **105**, 570–580 (2013).
17. K. Wilson *et al.*, *Nat. Commun.* **4**, 2896 (2013).
18. H. W. Yang, S. R. Collins, T. Meyer, *Nat. Cell Biol.* **18**, 191–201 (2016).
19. N. Komatsu *et al.*, *Mol. Biol. Cell* **22**, 4647–4656 (2011).
20. T. Inoue, W. D. Heo, J. S. Grimley, T. J. Wandless, T. Meyer, *Nat. Methods* **2**, 415–418 (2005).
21. C. Gabella *et al.*, *Curr. Biol.* **24**, 1126–1132 (2014).
22. X. Yao, L. Cheng, J. G. Forte, *J. Biol. Chem.* **271**, 7224–7229 (1996).
23. V. Delorme *et al.*, *Dev. Cell* **13**, 646–662 (2007).
24. M. Oser, J. Condeelis, *J. Cell. Biochem.* **108**, 1252–1262 (2009).
25. Z. Shi, Z. T. Graber, T. Baumgart, H. A. Stone, A. E. Cohen, *Cell* **175**, 1769–1779.e13 (2018).
26. M. Bergert, S. D. Chandradoss, R. A. Desai, E. Paluch, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 14434–14439 (2012).
27. T. Das *et al.*, *Nat. Cell Biol.* **17**, 276–287 (2015).
28. P. K. Mattila, F. D. Batista, B. Treanor, *J. Cell Biol.* **212**, 267–280 (2016).
29. A. Bisaria, A. Hayer, D. Garbett, D. Cohen, T. Meyer, Code for: Membrane-proximal F-actin restricts local membrane protrusions and directs cell migration, Zenodo (2020); https://doi.org/10.5281/zenodo.3734171.

ACKNOWLEDGMENTS

We thank members of the Meyer laboratory for helpful comments and discussions, Y. Fan and N. Ratnayeke for help with developing code and analysis methods, and O. Davydenko for help with the Leica total internal reflection fluorescence (TIRF) system. **Funding:** This work was supported by grants GM127026 and S10OD018073 to T.M., an NSF Graduate Research Fellowship to A.B., grant GM116328 to D.G., and a Stanford Center for Systems Biology Seed Grant to A.H. and D.C. **Author contributions:** A.B. and T.M. conceptualized the study and methodology, acquired funding, and wrote the manuscript. A.B. performed the investigation, validation, visualization, data curation, and data analysis. A.B. and A.H. developed the software. A.B., D.G., A.H., and D.C. developed resources for the study. D.G., A.B., A.H., and T.M. reviewed and edited the manuscript. T.M. supervised the study. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data are available in the supplementary materials. Code is available on Zenodo (*29*). Raw time-lapse images will be provided upon request.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1205/suppl/DC1
Materials and Methods
Figs. S1 to S9
References (*30*–*43*)
Movies S1 to S3

17 July 2019; resubmitted 24 February 2020
Accepted 9 April 2020
10.1126/science.aay7794

STRUCTURAL BIOLOGY

Structures of cell wall arabinosyltransferases with the anti-tuberculosis drug ethambutol

Lu Zhang[1,2]*, Yao Zhao[1,3,4]*, Yan Gao[5], Lijie Wu[1], Ruogu Gao[4,6], Qi Zhang[1], Yinan Wang[1,4], Chengyao Wu[1], Fangyu Wu[2], Sudagar S. Gurcha[7], Natacha Veerapen[7], Sarah M. Batt[7], Wei Zhao[2], Ling Qin[1]†, Xiuna Yang[1], Manfu Wang[1], Yan Zhu[1], Bing Zhang[1], Lijun Bi[6], Xian'en Zhang[6], Haitao Yang[1], Luke W. Guddat[8], Wenqing Xu[1], Quan Wang[1,6]‡, Jun Li[1]‡, Gurdyal S. Besra[7]‡, Zihe Rao[1,2,5,6]‡

The arabinosyltransferases EmbA, EmbB, and EmbC are involved in *Mycobacterium tuberculosis* cell wall synthesis and are recognized as targets for the anti-tuberculosis drug ethambutol. In this study, we determined cryo–electron microscopy and x-ray crystal structures of mycobacterial EmbA-EmbB and EmbC-EmbC complexes in the presence of their glycosyl donor and acceptor substrates and with ethambutol. These structures show how the donor and acceptor substrates bind in the active site and how ethambutol inhibits arabinosyltransferases by binding to the same site as both substrates in EmbB and EmbC. Most drug-resistant mutations are located near the ethambutol binding site. Collectively, our work provides a structural basis for understanding the biochemical function and inhibition of arabinosyltransferases and the development of new anti-tuberculosis agents.

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*), is one of the oldest diseases known to infect humans but remains a major cause of morbidity and mortality, resulting in more than 1.5 million deaths each year (*1*). Ethambutol is one of the five first-line anti-TB drugs that are currently in clinical use to treat TB (*1*). Ethambutol is particularly effective in combination therapy against multidrug-resistant forms of this infectious disease (*1*). The membrane-embedded Emb proteins EmbA, EmbB, and EmbC, which are involved in cell wall biosynthesis, are regarded as the targets of ethambutol because mutations to these proteins result in TB strains that are clinically resistant to this drug. Most of these resistance sites occur within EmbB (*2*–*5*).

The cell wall of *Mtb* is more complex than those of Gram-negative and Gram-positive bacteria. Its core structure, mycolyl-arabinogalactan-peptidoglycan, is composed of three highly unusual elements covalently linked together: (i) long-chain mycolic acids, (ii) a highly branched arabinogalactan (AG), and (iii) a cross-linked network of peptidoglycans. The other key component of the cell wall, lipoarabinomannan (LAM), is a phosphatidyl-*myo*-inositol–derived glycolipid that contains mannan and arabinan domains. As a virulence factor, LAM plays a key role in host–pathogen interactions, as well as in modulating the host immune response during infection (*6*–*10*). The composition of AG and LAM suggests that at least seven different arabinosyltransferases (AraTs) are involved in the assembly of the arabinan domain (*6*). Because the Emb proteins belong to the same family of AraTs, their amino acid sequences are highly similar (sharing ~40% identity). Both EmbA and EmbB have important roles in the formation of the α(1→3) linkage on the terminal hexaarabinofuranosyl motif of AG (*11*) (Fig. 1A). Furthermore, these proteins are suggested to function in a coordinating way by forming a heterodimer within cells (*12*), whose reaction product is further catalyzed by AftB by forming the terminal β(1→2) linkage at the AG nonreducing end (*13*) (Fig. 1A). The product of the reactions catalyzed by EmbA and EmbB and by AftB serves as the mycolic acid attachment site for AG (*14*). EmbC functions by forming α(1→5) glycosidic linkages, leading to the linear elongation of the arabinan chain of LAM (Fig. 1B) (*15*). Decaprenyl-phosphate-arabinose (DPA) is the only known arabinose donor for these AraTs in mycobacterial species (*16*). In *Mtb*, the *embA*, *embB*, and *embC* genes are clustered in the genome and have all been shown to be essential for in vitro growth of strain H37Rv by analysis of saturated Himar1 transposon libraries (*17*). *embA* and *embC* have also been shown to be essential in *Mtb* under normal growth conditions through the generation of deletion mutants (*18*, *19*).

The Emb proteins belong to the glycosyltransferase C (GT-C) superfamily (*15*, *20*). Despite their importance in cell wall synthesis, the three-dimensional (3D) structure for any component of an Emb protein is yet to be determined, with the exception of the C-terminal soluble domain of EmbC (*21*). Thus, it remains poorly understood how these enzymes function, how ethambutol exerts its mode of action, and how *emb* mutations lead to ethambutol resistance. To provide a foundation for understanding these phenomena, we have determined the 3D structures of *Mtb* and *Mycobacterium smegmatis* (*Msm*) EmbA-EmbB and *Msm* EmbC-EmbC ($EmbC_2$), with each in complex with ethambutol. Ethambutol is observed bound in the conserved active sites of both EmbB and EmbC. In addition, we have also elucidated the 3D structures of *Msm* EmbA-EmbB and $EmbC_2$ in complex with the donor and acceptor substrates, DPA and diarabinose, respectively. Unexpectedly, these structures show that an acyl carrier protein, AcpM, is bound to the cytoplasmic surface of each Emb protomer. Our studies provide molecular insights into how arabinose is transferred by these enzymes and how ethambutol binds at a location that overlaps with the region where the donor and acceptor substrates bind.

Characterization and structure determination of the EmbA-EmbB and $EmbC_2$ complexes

The *Mtb* and *Msm* EmbA-EmbB complexes and the *Msm* $EmbC_2$ complex were expressed in *Msm* cultures and then purified to homogeneity (figs. S1A and S2E). Unexpectedly, we found that the endogenous *Msm* acyl carrier protein AcpM (MSMEG_4326, 99 amino acids, ~11 kDa), which is involved in the biosynthesis of mycolic acids (*6*, *22*), is copurified with each of these samples by SDS–polyacrylamide gel electrophoresis (PAGE) silver staining and mass spectrometry (figs. S1, C and D, and S2, B and J). SDS-PAGE showed that EmbA and EmbB have a 1:1 stoichiometry in EmbA-EmbB complexes (fig. S1B), whose composition was further confirmed by mass spectrometry (fig. S1, E to H). *Msm* EmbC is shown as a single band of ~120 kDa by SDS-PAGE (fig. S2A) and a single band between 242 and 480 kDa by blue native (BN)–PAGE (fig. S2F). Taking into account the fact that detergents wrap around this complex and the existence of AcpM, the molecular weight determined by BN-PAGE is consistent with dimer formation of *Msm* EmbC observed from our solved structures (see below).

All EmbA-EmbB and $EmbC_2$ samples were shown to be active in cell-free activity assays. The established AraT activity assay of purified *Mtb* and *Msm* EmbA-EmbB complexes (Fig. 1C and fig. S1J), together with 2D heteronuclear single quantum correlation nuclear magnetic resonance analysis of EmbA and EmbB deletion

[1]Shanghai Institute for Advanced Immunochemical Studies, iHuman Institute, School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China. [2]State Key Laboratory of Medicinal Chemical Biology, Frontiers Science Center for Cell Response, College of Life Sciences, College of Pharmacy, Nankai University, Tianjin 300353, China. [3]CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, Shanghai 200031, China. [4]University of Chinese Academy of Sciences, Beijing 100101, China. [5]Laboratory of Structural Biology, Tsinghua University, Beijing 100084, China. [6]National Laboratory of Biomacromolecules and Key Laboratory of RNA Biology, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, CAS, Beijing 100101, China. [7]Institute of Microbiology and Infection, School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. [8]School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, QLD 4072, Australia.
*These authors contributed equally to this work.
†Present address: Division of Structural Biology, University of Oxford, Oxford OX3 7BN, UK.
‡Corresponding author. Email: wangq@ibp.ac.cn (Q.W.); lijun1@shanghaitech.edu.cn (J.L.); g.besra@bham.ac.uk (G.S.B.); raozh@tsinghua.edu.cn (Z.R.)

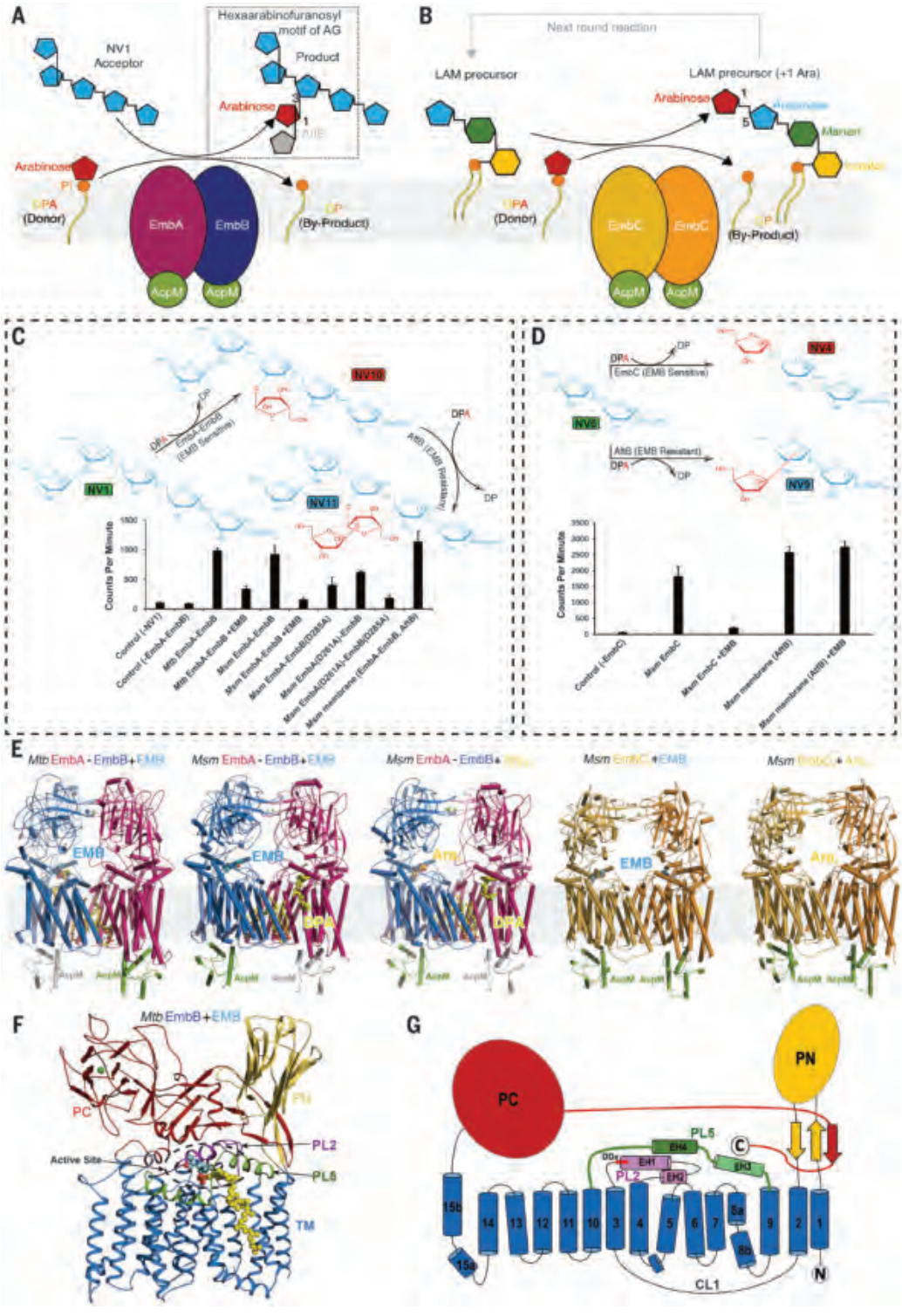


Fig. 1. Activity and overall structure of the EmbA-EmbB-AcpM$_2$ and EmbC$_2$-AcpM$_2$ complexes. (**A**) Schematic representation of the enzyme reaction catalyzed by the EmbA-EmbB-AcpM$_2$ complex, which transfers an arabinose residue from DPA to an arabinan acceptor (e.g., NV1) in an α(1→3) linkage. The extended product then serves as a precursor for subsequent extension by a β(1→2) AraT catalyzed by AftB, resulting in the synthesis of the terminal branching hexaarabinofuranosyl motif found in AG. (**B**) Schematic representation of the α(1→5) AraT reaction catalyzed by the EmbC$_2$-AcpM$_2$ complex, leading to elongation of the arabinan chain in the LAM precursor. The structure of the LAM precursor, which may contain several mannose and arabinose groups, is simplified here for clarity. (**C**) (Top) Designed reaction scheme illustrating α(1→3) AraT (EmbA-EmbB) and β(1→2) AraT (AftB) activity assays. (Bottom) Cell-free α(1→3) AraT activity of the purified WT EmbA-EmbB complexes and catalytic site mutations in the presence and absence of ethambutol (see also fig. S1J). NV1 was used as the acceptor and DP[^{14}C]A as the donor, as described in previous studies (*11*, *12*, *16*). The *Msm* membrane contains ethambutol-resistant AraT AftB. EMB, ethambutol. Data are mean values + SDs, calculated from three independent experiments. (**D**) (Top) Designed reaction scheme illustrating an α(1→5) AraT activity assay to characterize EmbC$_2$ activity. (Bottom) Cell-free α(1→5) AraT activity of the purified EmbC$_2$ complex in the presence and absence of ethambutol (see also fig. S2G). NV6 was used as the acceptor and DP[^{14}C]A as the donor, as described in previous studies (*13*, *16*). The *Msm* membrane contains ethambutol-resistant AraT AftB. Data are mean values + SDs, calculated from three independent experiments. (**E**) Overall view of cryo-EM structures of *Mtb* and *Msm* EmbA-EmbB-AcpM$_2$ complexes in complex with ethambutol or diarabinose, and cryo-EM and crystal structures of *Msm* EmbC$_2$-AcpM$_2$ complexes in complex with ethambutol or diarabinose. The AcpM protomer binds to each Emb protein in all complexes. Unmodeled AcpM protomer bound to *Mtb* EmbB or *Msm* EmbA is colored in gray. The drug EMB, the substrates DPA and Ara$_2$, and the lipids are shown as spheres. Ara$_2$, diarabinose. (**F**) The overall fold of Emb proteins represented by *Mtb* EmbB. The PN, PC, and TM domains are colored differently. The functionally important PL2 and PL5 are also highlighted. The location of active site is marked by a dashed circle. (**G**) Topological diagram of Emb proteins colored as in (F); the DDX motif is shown as three red spheres. EH, extracellular helix.

mutants (*23*) (fig. S1K), confirmed that they function as AraTs, which catalyze the formation of an α(1→3)-arabinofuranosyl linkage, in accordance with previous studies (*11*, *12*, *16*). The *Msm* $EmbC_2$ sample was confirmed to catalyze the formation of an α(1→5)-arabinofuranosyl linkage (*16*) (Fig. 1D and fig. S2G). More importantly, the cell-free AraT activity of both EmbA-EmbB and $EmbC_2$ is inhibited by ethambutol, confirming that ethambutol targets these proteins (Fig. 1, C and D, and figs. S1J and S2G).

Single-particle cryo–electron microscopy (cryo-EM) was used to determine four structures of Emb proteins—*Mtb* EmbA-EmbB-$AcpM_2$ bound with ethambutol, *Msm* EmbA-EmbB-$AcpM_2$ bound with ethambutol, *Msm* EmbA-EmbB-$AcpM_2$ bound with diarabinose, and *Msm* $EmbC_2$-$AcpM_2$ bound with ethambutol—at a resolution of 2.81 to 3.10 Å (Fig. 1E, figs. S3 to S7, and table S1). For the Emb protomers, most regions of the polypeptide for EmbA, EmbB, and EmbC could be traced, except for the cytoplasmic loop CL1 (residues 248 to 268) in *Mtb* EmbB and a periplasmic segment (residues 780 to 810) of *Msm* EmbC. The density for AcpM bound to the *Msm* EmbB or *Mtb* EmbA surface has a resolution of 3 to 5 Å, whereas the map for AcpM bound to the *Msm* EmbA or *Mtb* EmbB protomer was less clear and was not built (Fig. 1E and figs. S3F, S4F, and S5F).

For the crystallographic study of EmbC, homologs from *Mtb*, *Msm*, *M. marinum*, and *M. xenopi* were all expressed and purified (fig. S2A). However, only the crystal structure of *Msm* EmbC in complex with diarabinose could be determined at 3.3-Å resolution (Fig. 1E, fig. S8, and table S2). On the basis of the electron density, the structure of two molecules of AcpM (residues 3 to 86) could be built (Fig. 1E and fig. S8A).

Overall structures of the EmbA-EmbB-$AcpM_2$ and $EmbC_2$-$AcpM_2$ complexes

In our cryo-EM and crystal structures, EmbA and EmbB form a heterodimeric complex, whereas EmbC is a symmetric homodimer (Fig. 1E). In the EmbC dimer, two EmbC protomers are nearly identical and could well superimpose on each other (fig. S9A). The mode of dimerization is similar in the two complexes and is achieved by forming hydrophobic clusters between transmembrane (TM) domains close to the cytoplasmic and periplasmic sides (fig. S10, A and B). The individual EmbA, EmbB, and EmbC proteins all have a similar fold (Fig. 1, F and G, and fig. S10C) and contain common features such as a 15-helix TM domain, N- and C-terminal periplasmic domains (PDs) identified as PN and PC (Fig. 1, F and G, and fig. S9, D, E, and G), and periplasmic and cytoplasmic loops connecting the TM helices, some of which are structurally and physiologically important. Periplasmic loop 2 (PL2) contains two crossed helices (EH1 and EH2) and harbors the highly conserved catalytically relevant DDX motif (*24*) (Fig. 1, F and G, and fig. S13A); PL5 forms two tandem helices (EH3 and EH4) and contributes to the gating of DPA and dimerization of the complex (Figs. 1, F and G, and fig. S10B). The positively charged CL1 forms extensive interactions with AcpM (Fig. 1G and fig. S11).

In each structure, the PN domain (also PL1, which links TM1 and TM2) adopts a jelly-roll fold (fig. S9D), which is typical for polysaccharide binding units such as plant lectins or where carbohydrates act as enzyme substrates (*25*). The PC domain can be divided into two subdomains, with subdomain I displaying a mixed α and β structure and subdomain II exhibiting a jelly-roll fold (fig. S9E). This fold of the PC domain is similar to the previously reported C-terminal structure of the *Mtb* EmbC [Protein Data Bank (PDB) ID: 3PTY], as both contain a bound Ca^{2+} that is responsible for structural stability (*21*) (fig. S9, C and F). The last 30 amino acids in the PC domain embrace the PN domain, thus contributing to the stabilization of the entire PDs (Fig. 1, F and G). The active site is located in a pocket at the junction between the TM domain and the PDs (Fig. 1F), composed of PL2 to PL6, helix α6 in the PC domain, and residue Trp^{965} in the PC domain (represented by *Msm* EmbC) (fig. S9H). Substrates (diarabinose analog and DPA) or ethambutol are observed in this site (Fig. 1E), providing crucial information for understanding catalysis, drug inhibition, and resistance.

By providing structural information that differs from that of other known glycosyltransferases, the structures of the Emb proteins substantially broaden our understanding of the GT-C superfamily. A structural comparison with *Archaeoglobus fulgidus* AglB (*26*), *Campylobacter lari* PglB (*27*), *Cupriavidus metallidurans* ArnT (*28*), yeast STT3 in OST complex (*29*), yeast PMT1 in PMT1-PMT2 complex (*30*), and yeast ALG6 (*31*) shows that the Emb proteins possess periplasmic architectures that are distinctive in size and shape from those of other GT-C proteins (fig. S12). Features common to those of other GT-C members include (i) TM regions that share a common core of 11 TM helices with a similar fold (fig. S12) and (ii) crossed helices that resemble EH1 and EH2 in Emb proteins, bearing the conserved catalytically relevant D[N]D[E]X motif (fig. S12).

The *Msm* AcpM binds to the cytoplasmic face of both the EmbA-EmbB and $EmbC_2$ complexes, thus forming the EmbA-EmbB-$AcpM_2$ and $EmbC_2$-$AcpM_2$ complexes, respectively (Fig. 1E). It has a four-helix topology arranged in a right-handed bundle (fig. S11, A and H), which is similar to that of *Mtb* AcpM (PDB ID: 1KLP) (*32*) (fig. S9B). The AcpMs bind to each Emb protomer through extensive electrostatic interactions (fig. S11, B, C, and I). This type of assembly has not been observed in any other glycosyltransferase. The binding mode of AcpM to EmbA-EmbB and to $EmbC_2$ is similar, whereby helix α2 of AcpM and the connecting loops at its N and C terminus are intimately engaged with the CLs of the Emb proteins (fig. S11, A and H). This is consistent with the known role of α2 in AcpM as a contact site with its target proteins [e.g., AcpS (*33*)]. 4′-Phosphopantetheine (Ppant) is also observed covalently attached to the conserved Ser^{41}_{AcpM} located on α2 of AcpM and inserts into the gap between TM6 and TM7 and CL1 of the Emb protein in the cryo-EM structures of both the EmbA-EmbB and $EmbC_2$ complexes (fig. S11, E, H, and I). By contrast, in the crystal structure of EmbC, the side chain of Ser^{41} of AcpM interacts with the main chain of Arg^{247} of EmbC on the CL1 with no Ppant observed (fig. S11D). Mutagenesis and functional studies showed that when the interactions between EmbC and AcpM are disrupted, the produced LAM species become smaller (except for R352A) (fig. S11G), although the formation of the $EmbC_2$-$AcpM_2$ complex with these mutants and cell-free AraT activity are largely preserved (figs. S2, C and D, and S11F). A likely explanation is that the mode of AcpM binding changes in the complex, in turn affecting the ability of AcpM to modulate LAM synthesis in vivo. Sequence alignment demonstrates that this region is not conserved across species (fig. S13A), indicating the variable binding abilities within the different Emb proteins. Because AcpM plays a notable role in the biosynthesis of fatty acids and cell walls (*34*), the association with proteins involved in AG assembly or LAM synthesis could imply some clues for the physiological function of this pattern. Further investigation into the exact function of AcpM in the Emb complexes is needed.

Several functionally important lipids are also observed in these cryo-EM structures. The substrate DPA is observed endogenously bound to EmbA in the conserved donor binding cavity in the *Msm* EmbA-EmbB complexes (Figs. 1E and 2B and fig. S7E), and the reaction by-product DP (the leaving group of DPA) is observed in a similar position in EmbB in the *Mtb* EmbA-EmbB complex (Fig. 1E and fig. S7F). The identities of DPA and DP were subsequently confirmed by mass spectrometry (fig. S1I). In all EmbA-EmbB complexes, native cardiolipins were observed binding in a similar manner to the TM domain at the dimer interface (Fig. 1E and figs. S7G and S10A). Thus, these lipids play a key role in stabilizing the heterodimeric complex.

Substrate binding in the active sites of the EmbA-EmbB and $EmbC_2$ complexes

The Emb proteins have two substrates: an arabinose donor DPA and an acceptor arabinan.

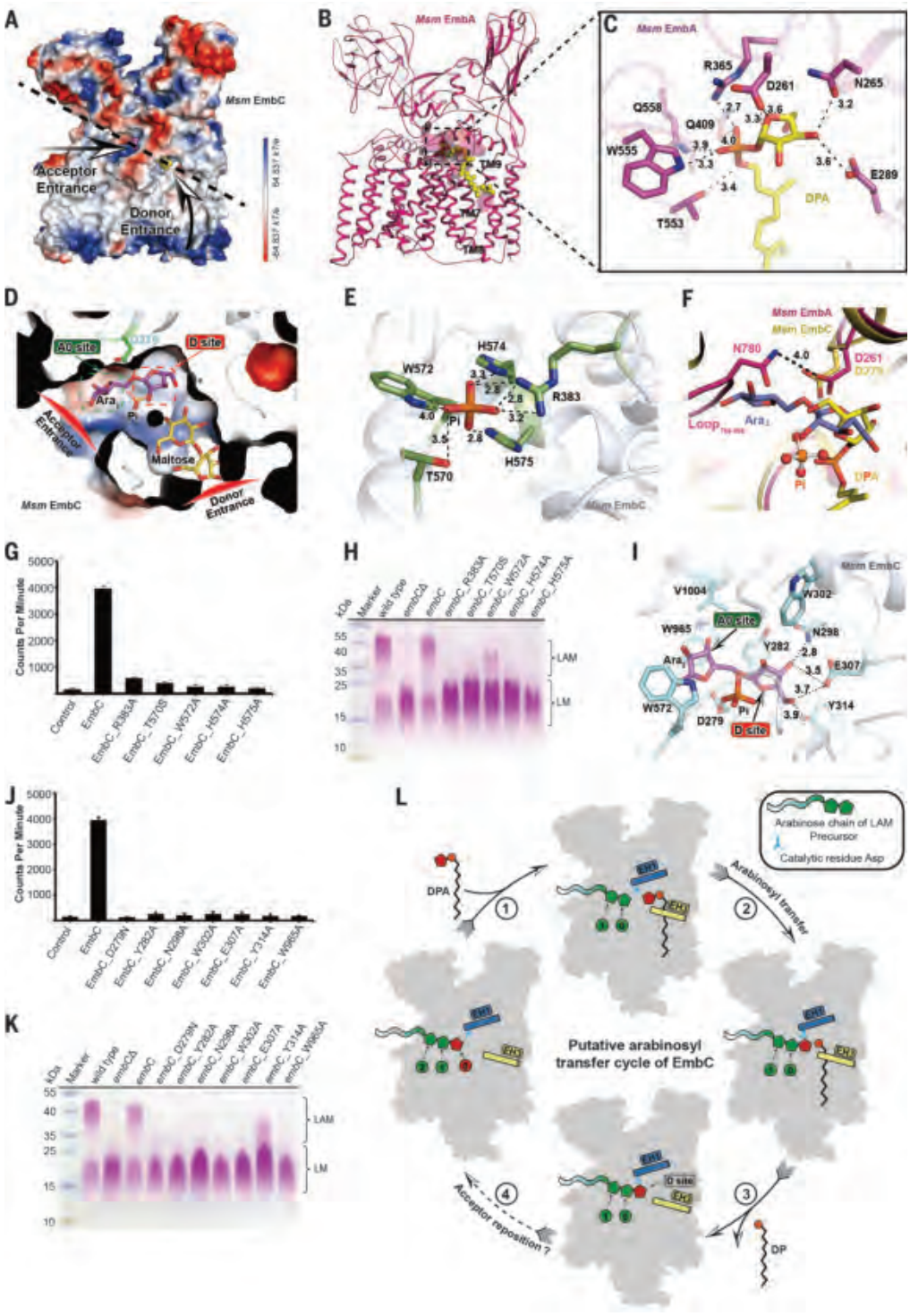


Fig. 2. Substrate binding in the active sites of the EmbA-EmbB and $EmbC_2$ complexes. (**A**) Overview of the substrate entrances to the active site in Emb proteins (represented by *Msm* EmbC, which is shown in electrostatic surface representation). The arrows indicate two entrances to the active site: one for the donor and one for the acceptor. The dashed line indicates the clipping position for (D). (**B**) Overview of *Msm* EmbA complexed with donor substrate DPA (which is shown as spheres). The head groups (arabinose and phosphate) bind in the active site through the donor entrance, whereas the tail groups bind to the TM region. (**C**) Zoomed-in view of DPA binding in the active site of *Msm* EmbA. Polar interactions are indicated by black dashed lines. The distance between the catalytic Asp^{261} and the C1 atom of DPA is marked by a pink dashed line. (**D**) Clipped view of the active site with acceptor entrance and donor entrance in *Msm* EmbC. Ara_2 (part of Ara_2OC8), maltose (part of detergent DDM), P_i (phosphate ion), and the catalytic residue (Asp^{279}) are shown as sticks. Ara_2 could mimic the terminal two arabinose groups of the reaction-elongated product by EmbC. The red and green boxes indicate the binding position (D site) for arabinose from the donor and the binding position (A_0 site) for arabinose from the terminal residue of the acceptor, respectively. (**E**) Phosphate binding site of *Msm* EmbC. P_i and surrounding residues are shown as sticks. Polar interactions are indicated by dashed lines. (**F**) Superposition between *Msm* EmbC (yellow) and *Msm* EmbA (magenta) in the active site. The D-site arabinose group of Ara_2 in *Msm* EmbC binds in the same position as that of DPA in *Msm* EmbA. The phosphate in *Msm* EmbC binds in a similar position as that of DPA in *Msm* EmbA. In *Msm* EmbA, the Asn^{780} on $loop_{766\text{-}806}$ interacts with the catalytic Asp^{261} and is close to the arabinose group of DPA. (**G**) Effect of mutated residues in the phosphate binding site on α(1→5) AraT activity of *Msm* $EmbC_2$ (*13*, *16*). Data are mean values + SDs, calculated from three independent experiments. (**H**) Effect of mutated residues in the phosphate binding site of *Msm* EmbC on LAM synthesis. LAM was extracted from recombinant *Msm* strains and analyzed by SDS-PAGE (on a denaturing, 16% acrylamide gel) and periodic acid–Schiff staining. *embC*Δ, *Msm embC* knockout mutant; *embC* or *embC*_R383A, *Msm embC* knockout mutant complemented with plasmid carrying *embC* WT or R383A mutant allele. (**I**) Close-up view of the Ara_2 binding site of *Msm* EmbC. Ara_2 and interacting residues are shown as sticks. Polar interactions are indicated by dashed lines. (**J**) Effect of mutated residues in the Ara_2 binding site on α(1→5) AraT activity of *Msm* $EmbC_2$ (*13*, *16*). Data are mean values +SDs, calculated from three independent experiments. (**K**) Effect of mutated residues in the Ara_2 binding site of *Msm* EmbC on LAM synthesis. LAM was extracted from recombinant *Msm* strains and analyzed by SDS-PAGE (on a denaturing, 16% acrylamide gel) and periodic acid–Schiff staining. *embC*Δ, *Msm embC* knockout mutant; *embC* or *embC*_D279N, *Msm embC* knockout mutant complemented with plasmid carrying *embC* WT or D279N mutant allele. (**L**) Proposed mechanism of arabinose transfer and chain elongation for EmbC. Helices EH1 and EH3 are indicated as cylinders. The catalytic Asp residue (Asp^{279} in *Msm* EmbC) is represented as blue sticks. Arabinose and phosphate groups are shown as a pentagon and circle, respectively. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; X, unspecified; and Y, Tyr.

To understand how substrates bind, we determined the crystal structure of diarabinose-bound *Msm* $EmbC_2$ and the cryo-EM structure of the diarabinose-bound *Msm* EmbA-EmbB complex (Fig. 1E). These structures show that there are two substrate entrances (the donor and acceptor entrances) leading to the active site (Fig. 2A). In the cryo-EM structure of *Msm* EmbA-EmbB, an endogenous donor substrate DPA is observed in the EmbA subunit. Its head part, the arabinose moiety and the phosphate group, binds in the active site, whereas the long decaprenyl tail extends out through the donor entrance and fits in a hydrophobic cavity created by TM7 to TM9 (Fig. 2B). The arabinose moiety forms polar interactions with surrounding residues Asp^{261}, Asn^{265}, and Glu^{289} on PL2, whereas the phosphate group is stabilized by polar bonds to Arg^{365}, Gln^{409}, Thr^{553}, Trp^{555}, and Gln^{558} (Fig. 2C). Asp^{261}, the catalytic residue on the DDX motif conserved in the GT-C family (fig. S13, A and B), is 3.3 Å from the C1 atom of DPA (Fig. 2C). Note that periplasmic $loop_{766\text{-}806}$ (which harbors residues 766 to 806) hangs down from the PC domain of *Msm* EmbA and blocks the acceptor entrance (fig. S14C). The Asn^{780} on this loop inserts into the active site and interacts with the Asp^{261} and is also close to the arabinose group of DPA (Fig. 2F). In EmbB and EmbC, the corresponding loop is either shorter and folded on the PC domain (in EmbB) or flexible in the solvent (in EmbC) (fig. S14C).

Consistent with the above observations, in the crystal structure of diarabinose-bound *Msm* $EmbC_2$, an endogenous phosphate ion appears to be trapped in the active site by diarabinose (part of Ara_2OC8) and maltose (part of detergent DDM) (Fig. 2D and fig. S8, B and C). It is bound to a positively charged region that includes Arg^{383}, His^{574}, and His^{575} and is near Thr^{570}, Trp^{572}, and the catalytic Asp^{279} (Fig. 2E). We propose that this phosphate represents the phosphate group of DPA, as it is superimposable with the phosphate group of DPA in *Msm* EmbA (Fig. 2F and fig. S14B). To verify this proposal, we measured the dissociation constant (K_d) values of DPA for the wild-type (WT) and mutant *Msm* EmbC proteins. The K_d for WT EmbC is 3.0 μM. For the EmbC mutants, the affinity is greatly reduced to 122 and 137 μM for the H574A and H575A mutants, respectively, and is undetectable for the R383A and T570S mutants (table S3). Furthermore, enzymatic activity is completely lost for these phosphate binding site mutants (Fig. 2G). Consistent with this finding, the in vivo LAM synthesis was almost completely inhibited when the *embC* knockout strain was complemented with the corresponding mutant alleles (for the W572A mutant, the LAM species become smaller) (Fig. 2H). This phosphate binding site for DPA is thus crucial for EmbC function.

The diarabinoside group, which is identified as part of the substrate analog Ara_2OC8 that was added during crystallization, binds between Asp^{279} and the phosphate (Fig. 2D and fig. S8, B and C). For clarity, we denote the positions of the two arabinofuranose rings as D site (arabinose from donor) and A_0 site (arabinose from the terminal residue of acceptor), with the D site being deeper in the pocket and the A_0 site being closer to the acceptor entrance (Fig. 2D). The arabinofuranose in the D site is sandwiched between Tyr^{282} and the phosphate group. Its hydroxyl groups interact with Asn^{298}, Glu^{307}, and Tyr^{314} through hydrogen bonds (Fig. 2I). The A_0-site arabinofuranose is clamped by Trp^{572} and Trp^{965} (Fig. 2I). Additionally, Trp^{302} and Val^{1004} form van der Waals interactions with the D-site and A_0-site arabinofuranoses, respectively (Fig. 2I). The side chain of the catalytic Asp^{279} points toward the α(1→5) glycosidic bond between the two arabinofuranose groups (Fig. 2, D and I). The K_d value of Ara_2OC8 with *Msm* EmbC is 36.7 μM (fig. S2H). When these residues are mutated to Ala or Asn, the enzymatic activity of *Msm* EmbC is reduced significantly or cannot be detected (Fig. 2J), confirming that these residues are essential for the function of EmbC. Consistent with this observation, these mutations almost completely abolish production of LAM species in the *Msm* EmbC knockout strain complemented with the same alleles (for the Y314A mutant, the LAM species become smaller) (Fig. 2K). This is in accord with a previous report that showed that the D279A mutant could not produce LAM (*24*). Considering that the two arabinoside groups are located on different sites of the catalytic site Asp^{279}, the diarabinoside appears to represent the reaction product after Asp^{279} catalyzes the formation of the α(1→5) linkage between the two arabinofuranose rings. The arabinofuranose in the D site is most likely mimicking the newly added residue in the product, which is superimposable with the arabinose group of DPA in *Msm* EmbA (Fig. 2F) and thus originates from donor DPA, whereas the arabinofuranose in the A_0 site resembles an arabinose from the terminal residue in the acceptor.

On the basis of the above structural and functional data, an arabinose transfer mechanism that completes one cycle of elongation can be proposed for EmbC (Fig. 2L): (i) First, the terminal arabinofuranose of a LAM precursor from a previous round of reaction stays in the catalytic pocket and binds at the A_0 site. The hydrophilic head group of DPA inserts into the pocket from the donor entrance, and its phosphate group and the donor arabinofuranose bind at the phosphate binding site and the D site, respectively. (ii) Next, the catalytic Asp activated the transfer of donor arabinofuranose to the arabinofuranose residue in the A_0 site by forming an α(1→5) linkage. As a result, the donor arabinofuranose becomes the terminal residue of the elongated product, which binds to the D site. The diarabinose binding mode in the crystal structure of EmbC represents the terminal diarabinose of the elongated product. (iii) The by-product of the reaction, DP, leaves the catalytic pocket. (iv) In the final step, the elongated product is reloaded, or repositioned by an unknown mechanism, to the active site such that the terminal arabinofuranose again occupies the A_0 site, and the next elongation reaction can proceed.

In the cryo-EM map of the *Msm* EmbA-EmbB complex, the density of the diarabinose (part of the incubated tetra-arabinose) was found within the active site of EmbB subunit (fig. S14D), a location similar to the diarabinose observed in the crystal structure of *Msm* $EmbC_2$ (fig. S14C). The diarabinose bound in *Msm* EmbB results from interactions of Trp^{578}, His^{580}, Trp^{972}, and Trp^{1012} with the A_0-site group and of Tyr^{288}, Asn^{304}, Glu^{313}, and Arg^{495} with the D-site group; the catalytic residue Asp^{285} forms a polar interaction with the oxygen atom on the glycosidic bond (fig. S14D). These residues are conserved among EmbB and EmbC proteins (fig. S13A). Similar to the diarabinose in the crystal structure of *Msm* EmbC, the D-site arabinose group in *Msm* EmbB is thought to be provided by the donor DPA (i.e., the arabinose has been cleaved from the donor DPA) upon superimposition with DPA of *Msm* EmbA (fig. S14A). Thus, the A_0-site arabinose represents the sugar moiety from the acceptor substrate so that the disaccharide mimics the product in the reaction center. Given that the donor substrate DPA is observed in its pre-catalytic state in EmbA, it is plausible that the binding of substrates is a sequentially coupled process with DPA binding followed by acceptor binding. Furthermore, the likely one-step reaction catalyzed by the EmbA-EmbB complex suggests that the newly formed glycosidic bond occurs and is followed by release of all products. The next cycle must be initiated by the binding of another molecule of DPA. Future investigations will be needed to elucidate how the entire catalytic cycle works.

Structural basis for ethambutol inhibition on EmbB and EmbC

Our functional data confirm that ethambutol inhibits the enzymatic transferase activity of both the EmbA-EmbB and $EmbC_2$ complexes in vitro (Fig. 1, C and D, and figs. S1J and S2G). These results are in agreement with the observation that branching of the terminal hexa-arabinan motif in AG, as well as synthesis of LAM, can be inhibited by ethambutol (*23*, *35*). The K_d for ethambutol binding was measured to be 0.42 μM for *Mtb* EmbA-EmbB (Fig. 3B and table S3) and 0.31 μM for *Msm* EmbA-EmbB (Fig. 4C and table S3); thus, strong binding is observed in both cases. For

Msm EmbC, the K_d value was measured to be 11.1 μM (fig. S2I and table S3), indicating a relatively weaker binding affinity for this class of Emb protein.

To elucidate how ethambutol binds, cryo-EM structures of the *Mtb* and *Msm* EmbA-EmbB complexes, as well as *Msm* $EmbC_2$, all in complex with ethambutol, were determined at resolutions of 2.97, 2.90, and 2.81 Å, respectively (Fig. 1E). Analysis of these three maps showed that density consistent with that of ethambutol is located within the active site of the EmbB and EmbC subunits (Figs. 3, A and C, and fig. S15D). By contrast, no density for ethambutol was observed in the EmbA subunit. However, there is density for the endogenous donor substrate DPA, whose hydrophilic moiety bound to *Msm* EmbA is in a similar location as ethambutol when bound to EmbB (Fig. 3, F and G). We thus suggest that ethambutol preferentially binds to EmbB and EmbC rather than to EmbA, a conclusion in line with clinical drug resistance studies (*3*–*5*, *36*–*40*). Considering the similarity in ethambutol binding in all determined complexes (fig. S15F), we focused our analysis of its binding mode on the basis of the *Mtb* EmbA-EmbB complex. In this case, the two charged imino groups of ethambutol, NH1 and NH2, play key roles in binding. They form three electrostatic

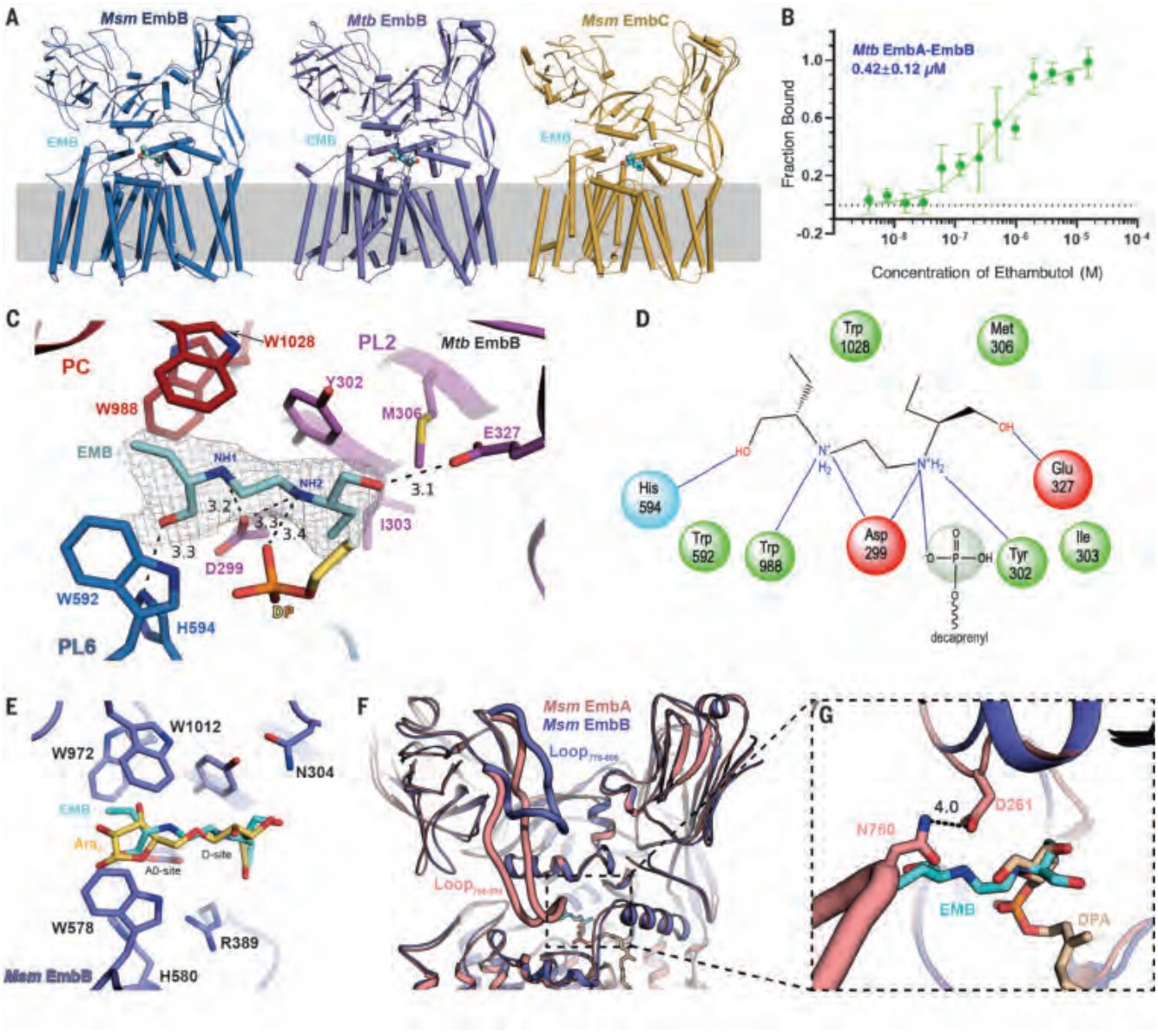


Fig. 3. Structural basis for ethambutol inhibition of EmbB and EmbC. (**A**) Cartoon representations of ethambutol-bound *Msm* EmbB (left), *Mtb* EmbB (middle), and *Msm* EmbC (right). Ethambutol is shown as cyan spheres. (**B**) Binding affinity of ethambutol with the *Mtb* EmbA-EmbB complex measured by the microscale thermophoresis (MST) assay. The K_d value is provided. Data are representative mean values + SDs, calculated from three independent experiments. (**C**) Structural details of ethambutol binding to *Mtb* EmbB. PL2, PL6, and the PC domain are colored in purple, blue, and red, respectively. Interacting residues are shown as sticks. Polar interactions are indicated by dashed lines. The cryo-EM map density for ethambutol (threshold 0.4) is shown as gray mesh. (**D**) Schematic diagram of the interaction between ethambutol and *Mtb* EmbB. (**E**) Superposition of the active site region of the ethambutol (cyan)–bound *Msm* EmbB to the disaccharide (yellow)–bound *Msm* EmbB. (**F**) Superposition of *Msm* EmbA (pink) to *Msm* EmbB (blue) on their periplasmic regions. $Loop_{766\text{-}806}$ in *Msm* EmbA, which is longer, blocks the acceptor entrance to the active site, whereas the corresponding $loop_{776\text{-}806}$ in *Msm* EmbB is folded inside the PC domain. (**G**) Zoomed-in view of the active site upon superposition in (F). Ethambutol (cyan) in *Msm* EmbB clashes with the arabinose group in DPA and Asn^{780} in *Msm* EmbA.

interactions with EmbB (the distances of NH1-Asp^{299}, NH2-Asp^{299}, and NH2-DP are 3.2, 3.3, and 3.4 Å, respectively) by positioning themselves between Asp^{299} and the phosphate group of the DP (Fig. 3, C and D). In addition, π-cation interactions are formed by NH1 with Trp^{988} and NH2 with Tyr^{302} (Fig. 3, C and D). The two hydroxyl groups of ethambutol form hydrogen bonds with His^{594} and Glu^{327}, respectively, whereas the two hydroxybutanyl groups form van der Waals interactions with Ile^{303} and Met^{306} on PL2, Trp^{592} on PL6, and Trp^{1028} on the PC domain (Fig. 3, C and D).

When superimposing the EmbB (or EmbC) subunits in complex with ethambutol and

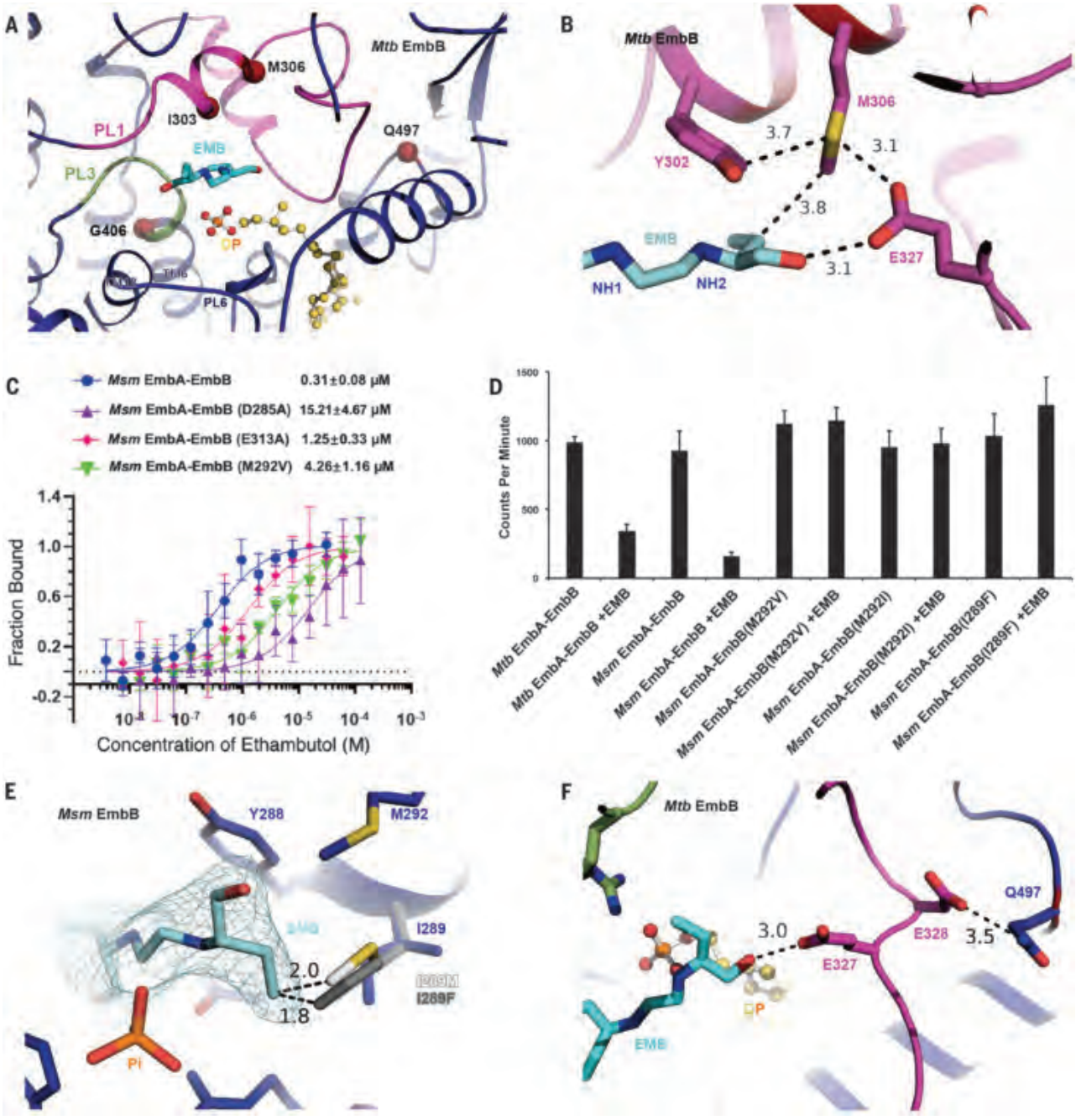


Fig. 4. Structural interpretation for ethambutol resistance. (**A**) Cartoon representation of the ethambutol binding pocket in *Mtb* EmbB. Ethambutol is shown as sticks. Residue sites bearing the drug-resistant hotspot mutants are denoted with dark red spheres. (**B**) Structural details of the Met^{306} interaction environment in *Mtb* EmbB. Dashed lines indicate the distance of surrounding residues or ethambutol to Met^{306}. (**C**) Binding affinity of ethambutol to *Msm* EmbA-EmbB or its ethambutol-resistance–associated mutants equivalent to that of *Mtb* EmbB, measured by the MST assay. The K_d values are provided; data are representative mean values + SDs, calculated from three independent experiments. (**D**) Cell-free α(1→3) AraT activity of the ethambutol-resistant mutations on EmbB (*Msm* M292V, M292I, and I289F) in the absence and presence of ethambutol (*11*, *12*, *16*). The WT *Msm* and *Mtb* EmbA-EmbB complexes were used as controls. Data are mean values + SDs, calculated from three independent experiments. (**E**) Mutations I289M (pale gray) and I289F (dark gray) affect ethambutol binding by forming steric clashes that ultimately lead to ethambutol resistance. The cryo-EM map for ethambutol (threshold 0.3) is shown as cyan mesh. (**F**) Drug-resistant hotspot Gln^{497}, surrounding the drug binding pocket.

disaccharide, we found that they overlapped with a high degree of similarity (Fig. 3E and fig. S15E). Because the disaccharide represents the arabinose group from both donor and acceptor, we infer that ethambutol inhibits the arabinose transfer reaction by competing with the binding of both substrates in the active site. This interpretation is consistent with the hypothesis that ethambutol interferes with arabinose transfer, as evidenced by the rapid accumulation of DPA in ethambutol-treated *Msm* cells (*41*).

Superimposing the structure of *Msm* EmbA onto *Msm* EmbB reveals that the conformation of loop$_{766\text{-}806}$ of *Msm* EmbA not only locks DPA in the active site but also hinders ethambutol from binding in the active site (Fig. 3, F and G). Considering the structural similarity between *Msm* and *Mtb* and the sequence homology (fig. S13A) among mycobacterial EmbA proteins, this long loop is identified as a conserved and distinctive feature of EmbA that traps the substrate and prevents ethambutol from targeting it.

Structural interpretation for ethambutol resistance

Numerous mutations in the *embCAB* locus have been identified by detection of increased minimal inhibitory concentration (MIC) in ethambutol-resistant *Mtb* strains compared with the WT strain. These mutations, most of which are changes to the *embB* gene, are likely the result of stress from ethambutol exposure. Mutations to three sites (Met^{306}, Gly^{406}, and Gln^{497}) in *Mtb* EmbB (*3*–*5*, *36*) (Fig. 4A) are regarded as resistance hotspots. Met^{306}, which is conserved in all EmbB proteins (fig. S13A), is the most frequently observed mutation site (*2*, *3*, *42*). In the ethambutol-bound *Mtb* EmbB structure, the side chain of Met^{306} is directly involved in ethambutol binding through van der Waals contacts with the hydroxybutanyl group on NH2 of ethambutol (Fig. 4B). Any changes in this site could affect drug binding. Drug-resistant isolates have favored mutations such as M306V and M306I to *Mtb* EmbB (*3*, *4*, *35*). Mutation of the equivalent methionine in *Msm* EmbB (Met^{292}) to Val results in a ~13-fold decrease in binding affinity to ethambutol (Fig. 4C), whereas AraT activity of the mutants (equivalent to *Msm* M292V and M292I) remains unaffected and resistant to ethambutol (Fig. 4D). Met^{306} is also involved in nonpolar interaction with the surrounding residues Tyr^{302} and Glu^{327} (Fig. 4B), the latter of which also interacts with ethambutol (Fig. 3C). Mutations on Met^{306} would therefore likely also change the interaction network to affect ethambutol binding. Two other mutant hotspot residues, Gly^{406} and Gln^{497}, are both more than 10 Å from ethambutol and thus have no direct interaction with the drug. Mutation of Gln^{497} will no longer permit an interaction with Glu^{328} and, as a result, may lead to a disruption of the Glu^{327}-ethambutol interaction (Fig. 4F). It is unclear how mutation to Gly^{406} leads to drug resistance, given its location at the junction of PL3 and TM6 (Fig. 4A), but it is likely that a bulkier side-chain mutation could result in steric hindrance that transmits to conformational change at the ethambutol binding site. Mutations on Ile^{289} of *Msm* EmbB (I289M and I289F), which is conserved in ethambutol-sensitive mycobacterial organisms [*Msm* and *M. bovis* (*43*)] (fig. S13A), are also known to be responsible for ethambutol resistance in *Msm* (*44*). Structural analysis shows that the I289M and I289F mutations would sterically hinder ethambutol binding but not affect enzymatic activity (Fig. 4, D and E), therefore resulting in resistance.

To further investigate the resistance to EmbB, 1814 resistant sites from 61 studies documented in the MycoResistance database (www.hmulinglab.org/MycoResistance/) (*42*) were manually selected, with the top 16 mutation sites mapped onto the *Mtb* EmbB (fig. S15A). Of these resistance mutations (including the hotspots mentioned above), ~73.7% occur in residues on PL2, which has proven pivotal to both drug binding and catalysis, further emphasizing the importance of PL2 as a hotspot for the binding of new drug candidates. PL3 and PL5 harbor 8.7 and 8.1% of mutations, respectively, with the remainder found on TM4, CL2, and TM5 (fig. S15B). Among the mutation sites on PL2, four D299E mutated strains were found to be associated with ethambutol resistance. In some glycosyltransferases, it has been reported that the corresponding Glu can serve as the catalytic residue (*20*, *45*–*47*), which indicates that this change can occur and the longer side chain can simultaneously continue to act as a catalytic nucleophile.

The clinically relevant ethambutol-resistance mutation sites in *Mtb* EmbC (*3*, *4*, *36*–*40*) are also highly conserved (fig. S15C) in *Msm* EmbC. Mapping of these sites onto the *Msm* $EmbC_2$ structure showed that they are close to the ethambutol binding site (fig. S15C), with substantial clustering occurring within PL2. Thus, the resistance data map well to the structure of *Msm* EmbC and are consistent with resistance analysis on *Mtb* EmbB.

Outlook

We have demonstrated that a heterodimeric EmbA-EmbB and a homodimeric EmbC form functional AraT complexes associated with two copies of AcpM and that ethambutol targets the active site of EmbB and EmbC. The DPA- and diarabinose-bound structures of the EmbA-EmbB and $EmbC_2$ complexes allow us to understand the structural features required for catalysis. Contrastingly, the EmbA-EmbB complex catalyzes a branching reaction, whereas $EmbC_2$ catalyzes a multicircle elongation reaction (Fig. 1, A and B). On the basis of these structures and supporting data, we propose that ethambutol functions by competing with the substrates for binding to the EmbB and EmbC subunits. Its binding mode almost precisely overlaps with the disaccharide product analog; as a result, *Mtb* may have limited opportunity to develop resistance without compromising its own ability to successfully construct its elaborate cell wall. This property is attractive because it is well known that coadministration of multiple drugs that have different targets is an effective way to slow development of resistance. Understanding how ethambutol interacts with other possible targets will enable the complete definition of ethambutol's mode of action. Because EmbB and EmbC and their orthologs are well conserved across mycobacteria, the development of drugs that are broadly effective against TB and other human pathogenic infectious diseases, including leprosy (caused by *Mycobacterium leprae*), is feasible.

REFERENCES AND NOTES

1. World Health Organization, "Global tuberculosis report 2019" (2019); www.who.int/tb/publications/global_report/en/.
2. H. Safi, B. Sayers, M. H. Hazbón, D. Alland, *Antimicrob. Agents Chemother.* **52**, 2027–2034 (2008).
3. Q. Sun *et al.*, *Antimicrob. Agents Chemother.* **62**, e01279-17 (2017).
4. L. L. Zhao *et al.*, *Antimicrob. Agents Chemother.* **59**, 2045–2050 (2015).
5. F. Brossier *et al.*, *Antimicrob. Agents Chemother.* **59**, 4800–4808 (2015).
6. M. Jankute, J. A. Cox, J. Harrison, G. S. Besra, *Annu. Rev. Microbiol.* **69**, 405–423 (2015).
7. I. Vergne, M. Gilleron, J. Nigou, *Front. Cell. Infect. Microbiol.* **4**, 187 (2015).
8. A. K. Mishra, N. N. Driessen, B. J. Appelmelk, G. S. Besra, *FEMS Microbiol. Rev.* **35**, 1126–1157 (2011).
9. V. Briken, S. A. Porcelli, G. S. Besra, L. Kremer, *Mol. Microbiol.* **53**, 391–403 (2004).
10. M. Gilleron, J. Nigou, D. Nicolle, V. Quesniaux, G. Puzo, *Chem. Biol.* **13**, 39–47 (2006).
11. V. E. Escuyer *et al.*, *J. Biol. Chem.* **276**, 48854–48862 (2001).
12. S. Khasnobis *et al.*, *Chem. Biol.* **13**, 787–795 (2006).
13. M. Seidel *et al.*, *J. Biol. Chem.* **282**, 14729–14740 (2007).
14. M. McNeil, M. Daffe, P. J. Brennan, *J. Biol. Chem.* **266**, 13217–13223 (1991).
15. S. Berg, D. Kaur, M. Jackson, P. J. Brennan, *Glycobiology* **17**, 35R–56R (2007).
16. R. E. Lee, P. J. Brennan, G. S. Besra, *Glycobiology* **7**, 1121–1128 (1997).
17. M. A. DeJesus *et al.*, *mBio* **8**, e02133-16 (2017).
18. A. G. Amin *et al.*, *Microbiology* **154**, 240–248 (2008).
19. R. Goude, A. G. Amin, D. Chatterjee, T. Parish, *J. Bacteriol.* **190**, 4335–4341 (2008).
20. L. L. Lairson, B. Henrissat, G. J. Davies, S. G. Withers, *Annu. Rev. Biochem.* **77**, 521–555 (2008).
21. L. J. Alderwick *et al.*, *PLOS Pathog.* **7**, e1001299 (2011).
22. M. L. Schaeffer, G. Agnihotri, H. Kallender, P. J. Brennan, J. T. Lonsdale, *Biochim. Biophys. Acta* **1532**, 67–78 (2001).
23. R. E. Lee, W. Li, D. Chatterjee, R. E. Lee, *Glycobiology* **15**, 139–151 (2005).
24. S. Berg *et al.*, *J. Biol. Chem.* **280**, 5651–5663 (2005).
25. A. B. Boraston, D. N. Bolam, H. J. Gilbert, G. J. Davies, *Biochem. J.* **382**, 769–781 (2004).
26. S. Matsumoto *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17868–17873 (2013).
27. C. Lizak, S. Gerber, S. Numao, M. Aebi, K. P. Locher, *Nature* **474**, 350–355 (2011).
28. V. I. Petrou *et al.*, *Science* **351**, 608–612 (2016).
29. R. Wild *et al.*, *Science* **359**, 545–550 (2018).
30. L. Bai, A. Kovach, Q. You, A. Kenny, H. Li, *Nat. Struct. Mol. Biol.* **26**, 704–711 (2019).
31. J. S. Bloch *et al.*, *Nature* **579**, 443–447 (2020).
32. H. C. Wong, G. Liu, Y.-M. Zhang, C. O. Rock, J. Zheng, *J. Biol. Chem.* **277**, 15874–15880 (2002).
33. K. D. Parris *et al.*, *Structure* **8**, 883–895 (2000).
34. L. Kremer *et al.*, *J. Biol. Chem.* **276**, 27967–27974 (2001).
35. B. Cuevas-Córdoba *et al.*, *Antimicrob. Agents Chemother.* **59**, 5455–5462 (2015).
36. C. Plinke *et al.*, *J. Antimicrob. Chemother.* **65**, 1359–1367 (2010).
37. Y. Xu, H. Jia, H. Huang, Z. Sun, Z. Zhang, *BioMed Res. Int.* **2015**, 951706 (2015).

38. S. V. Ramaswamy *et al.*, *Antimicrob. Agents Chemother.* **44**, 326–336 (2000).
39. W. L. Huang, T. L. Chi, M. H. Wu, R. Jou, *J. Clin. Microbiol.* **49**, 2502–2508 (2011).
40. A. Sandgren *et al.*, *PLOS Med.* **6**, e1000002 (2009).
41. B. A. Wolucka, M. R. McNeil, E. de Hoffmann, T. Chojnacki, P. J. Brennan, *J. Biol. Chem.* **269**, 23328–23335 (1994).
42. E. Dai *et al.*, *Database* **2019**, baz074 (2019).
43. *Tuberculosis* **88**, 102–105 (2008).
44. M. A. Lety, S. Nair, P. Berche, V. Escuyer, *Antimicrob. Agents Chemother.* **41**, 2629–2633 (1997).
45. H. Ihara *et al.*, *Glycobiology* **17**, 455–466 (2007).
46. T. Kubota *et al.*, *J. Mol. Biol.* **359**, 708–727 (2006).
47. T. A. Fritz, J. Raman, L. A. Tabak, *J. Biol. Chem.* **281**, 8613–8619 (2006).

ACKNOWLEDGMENTS

We thank S. Zhao, S. Wang, and M. Wu (ShanghaiTech University), Z. Du and H. Jin (Nankai University), and X. Cheng (Shanghai Institute of Materia Medica, Chinese Academy of Sciences) for assisting with computational and functional experiments; W. Chen, J. Chen, and W. Zhu (Analytical Chemistry Platform of Shanghai Institute for Advanced Immunochemical Studies) and C. Peng (Mass Spectrometry System of NCPSS) for technical support in mass spectrometry analysis; J. Bennett (University of Oxford) for the cryo-grid preparation; the National Centre for Protein Science Shanghai (protein expression and purification system) and The Molecular and Cell Biology Core Facility of the School of Life Science and Technology (ShanghaiTech University) for use of their instrumentation and technical assistance; X. Ye, D. Xiong, F. Xia, J. Bai, and Y. Wu (Peking University) and Y. Li (Tsinghua University) for help with carbohydrate chemistry; the Bio-Electron Microscopy Facility of ShanghaiTech University; Q. Sun for cryo-EM technical support; and the staff of beamlines BL17U1, BL18U1, and BL19U1 at Shanghai Synchrotron Radiation Facility (China), beamline BL41XU at SPring-8 (Japan), beamline I04-1 at Diamond Light Source (UK), and beamline PX at Swiss Light Source (Switzerland) for assistance during data collection. **Funding:** This work was supported by grants from the National Key R&D Program of China (2017YFC0840300) and Project of International Cooperation and Exchanges NSFC (81520108019) to Z.R., the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB29020000) to L.B., and the Medical Research Council (MR/S000542/1) to G.S.B. **Author contributions:** Z.R. conceived the project; Z.R., G.S.B., J.L., Q.W., L.Z., and Y.Zha. designed the experiments; Y.Zha., L.Z., Q.Z., Y.W., C.W., F.W. and L.Q. cloned and purified the Emb proteins and their mutants; Y.G., L.Z., Y.Zha., and Y.W. prepared cryo-EM samples; L.Z., Y.Zha., Y.G., R.G., L.Q., and Y.W. collected and processed the cryo-EM data; Y.G. and R.G. reconstructed the cryo-EM maps; L.Z., Y.Zha., R.G., and Y.G. built and refined the structure models; Y.Zha. and L.Z. grew the crystals and collected the diffraction data; J.L., L.W., and Y.Zha. solved the crystal structure of EmbC; S.S.G., N.V., and S.M.B. performed the enzymatic activity assay and synthesized chemical compounds for activity assay; L.B. and X.Z. provided the *emb* knockout *M. smegmatis* strain; and L.Z., Y.Zha., M.W., Y.Zhu., W.Z., Y.W., F.W., B.Z., and X.Y. performed mass spectrometry experiments and other biochemical experiments. All authors analyzed and discussed the results. L.Z., J.L., Q.W., Y.Zha., R.G., H.Y., L.W.G., W.X., G.S.B., and Z.R. prepared the manuscript, with help from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. The accession numbers for the 3D cryo-EM density maps reported in this paper are 30218 for *Mtb* EmbA-EmbB-AcpM_2 in complex with ethambutol, 30216 and 30219 for *Msm* EmbA-EmbB-AcpM_2 in complex with ethambutol and disaccharide, and 30217 for *Msm* EmbC_2-AcpM_2 in complex with ethambutol. The PDB IDs are 7BVC and 7BVG for the coordinates of the *Msm* EmbA-EmbB-AcpM_2 in complex with ethambutol and disaccharide, 7BVF for *Mtb* EmbA-EmbB-AcpM_2 in complex with ethambutol, and 7BVE and 7BVH for *Msm* EmbC_2-AcpM_2 in complex with ethambutol and diarabinose.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1211/suppl/DC1
Materials and Methods
Figs. S1 to S15
Tables S1 to S3
References (*48*–*62*)

15 January 2020; resubmitted 6 February 2020
Accepted 14 April 2020
Published online 23 April 2020
10.1126/science.aba9102

REPORTS

NANOPHOTONICS

Far-field excitation of single graphene plasmon cavities with ultracompressed mode volumes

Itai Epstein[1]*, David Alcaraz[1], Zhiqin Huang[2,3], Varun-Varma Pusapati[1], Jean-Paul Hugonin[4], Avinash Kumar[1], Xander M. Deputy[2,3], Tymofiy Khodkov[1], Tatiana G. Rappoport[5,6], Jin-Yong Hong[7], Nuno M. R. Peres[5,8], Jing Kong[7], David R. Smith[2,3], Frank H. L. Koppens[1,9]*

Acoustic graphene plasmons are highly confined electromagnetic modes carrying large momentum and low loss in the mid-infrared and terahertz spectra. However, until now they have been restricted to micrometer-scale areas, reducing their confinement potential by several orders of magnitude. Using a graphene-based magnetic resonator, we realized single, nanometer-scale acoustic graphene plasmon cavities, reaching mode volume confinement factors of ~5 × 10^{10}. Such a cavity acts as a mid-infrared nanoantenna, which is efficiently excited from the far field and is electrically tunable over an extremely large broadband spectrum. Our approach provides a platform for studying ultrastrong-coupling phenomena, such as chemical manipulation via vibrational strong coupling, as well as a path to efficient detectors and sensors operating in this long-wavelength spectral range.

Graphene plasmons (GPs) are propagating electromagnetic waves that are coupled to electron oscillations within a graphene sheet. As a result of their extraordinary properties of extreme confinement and low loss in the mid-infrared (MIR) to terahertz (THz) spectrum (*1*–*6*), they provide a platform to probe a variety of optical and electronic phenomena, including quantum nonlocal effects (*7*), molecular spectroscopy (*8*), and biosensing (*9*), and also enable access to forbidden transitions by bridging the scale between light and atoms (*10*). Furthermore, they offer a path toward miniaturizing optoelectronic devices in the long-wavelength spectrum, such as GP-based electro-optical detectors (*11*), modulators (*12*), and electrical excitation (*13*, *14*).

The confinement of GPs can be increased even further by placing the graphene sheet close to a metallic surface (*15*). Such a structure supports a highly confined asymmetric mode, which is referred to as an acoustic graphene plasmon (AGP) because of its linear energy versus momentum dispersion. When the graphene-metal distance is very small, AGPs are confined in-plane extensively to almost 1/300 of their equivalent free-space wavelength (*7*) and are vertically confined to the spacing between the metal and graphene (*16*). This confinement is accompanied by very little contribution from the metal to the AGP damping, even when including the quantum nonlocal effects of the metal (*17*).

This ability of AGPs to confine light to small dimensions is pivotal for strong light-matter interactions, especially in the MIR and THz spectra, where the wavelength is inherently large and limits the electromagnetic field confinement. The latter is eminently important at this spectral region, as it is where molecular resonances reside, and the detection of their spectral fingerprints is a requirement for potential applications in industrial processes, medicine, biotechnology, and security (*18*, *19*). Yet so far AGPs have only been observed over micrometer-scale areas, either as THz free-propagating waves (*7*, *15*) or in MIR grating couplers (*16*, *20*). Exploiting strong and ultrastrong light-matter interactions requires individual compact cavities for AGPs.

We realized single, nanometer-scale AGP cavities with ultrasmall mode volumes by the generation of a graphene plasmon magnetic resonance (GPMR), which enables the far-field excitation of AGP cavities over large areas without the need for lithographic patterning of either the surrounding environment or the graphene, and with no limitations on the light polarization. The GPMRs are formed by dispersing metallic nanocubes, with random locations and orientations, over a hexagonal boron nitride (hBN)/graphene van der Waals heterostructure.

[1]ICFO–Institut de Ciencies Fotoniques, Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain. [2]Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708, USA. [3]Center for Metamaterials and Integrated Plasmonics, Duke University, Durham, NC 27708, USA. [4]Université Paris-Saclay, Institut d'Optique Graduate School, CNRS, Laboratoire Charles Fabry, 91127 Palaiseau, France. [5]Centro de Física and Departamento de Física and QuantaLab, Universidade do Minho, P-4710-057 Braga, Portugal. [6]Instituto de Física–Universidade Federal do Rio de Janeiro, 21941-972 Rio de Janeiro RJ, Brazil. [7]Department of Electrical Engineering and Computer Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [8]International Iberian Nanotechnology Laboratory (INL), 4715-330 Braga, Portugal. [9]ICREA–Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain.
***Corresponding author. Email: itai.epstein@icfo.eu; frank.koppens@icfo.eu**

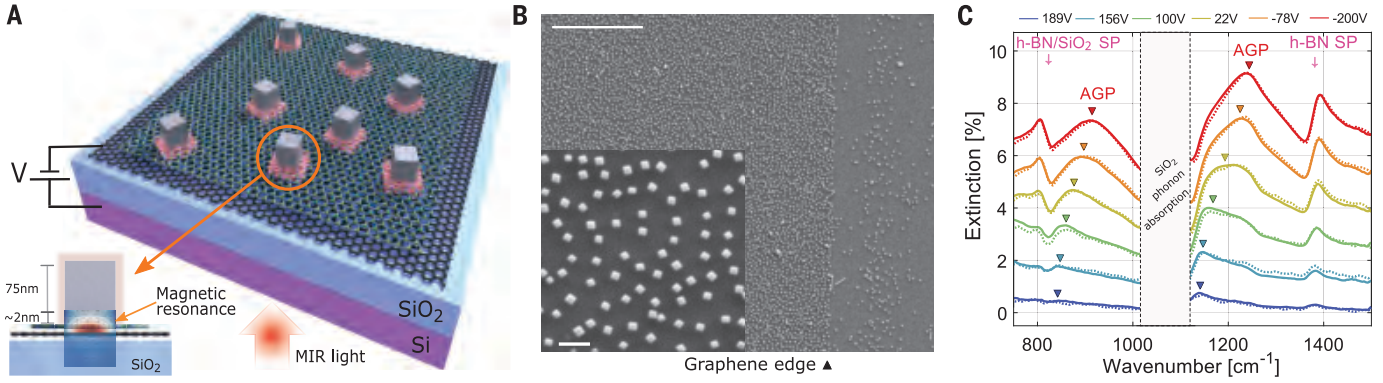


Fig. 1. Structure and optical response of a GPMR device. (**A**) Schematic of the GPMR device, which is composed of a Si/SiO_2 substrate, graphene/0.8 nm monolayer hBN heterostructure, and silver nanocubes. The inset shows the cross section of a single GPMR structure and the generated magnetic resonance between the nanocube and the graphene. (**B**) SEM images of the actual device, showing the graphene edge and deposited nanocubes. The inset presents a zoomed-in image of the nanocubes (scale bars, 5 μm and 200 nm, respectively). (**C**) A typical GPMR device extinction spectra measured in an FTIR apparatus, for different gate voltages (colors). Each colored triangle marks an AGP peak and its evolution with gate voltage. The downward arrows mark the location of the hBN and SiO_2 surface phonons. The solid and dotted lines correspond to the optical response for unpolarized and polarized light, respectively. The strong absorption band of the SiO_2 phonon is also marked.

In our GPMR device (Fig. 1A), nanometric silver nanocubes are randomly dispersed on top of a monolayer hBN/graphene heterostructure that is transferred onto a Si/SiO_2 substrate (*21*), which acts as a back gate for electrically doping the graphene. The nanocubes are randomly scattered over the graphene surface. The Fourier-transform infrared spectroscopy (FTIR) extinction spectra, $1 - [T(V)/T(V_{max})]$, measured from the device for different back-gate voltages (colors), are shown in Fig. 1C, where $T(V)$ is the transmission measured at a specific gate voltage V, and $T(V_{max})$ is transmission measured for the maximal voltage (corresponding to the lowest doping due to the intrinsic doping of the chemical vapor–deposited graphene). Two resonances can be observed in the gate-dependent extinction (marked with triangles), on both sides of the SiO_2 phonon absorption band (marked box). These resonances shift to higher frequencies with increasing graphene doping, which is consistent with the well-known behavior of plasmons in graphene (*16*, *22*, *23*). In addition, the supported surface phonons in the SiO_2 and hBN (marked with arrows) lead to the typical surface phonon–graphene plasmon hybridization in this type of heterostructure (*16*, *22*–*24*).

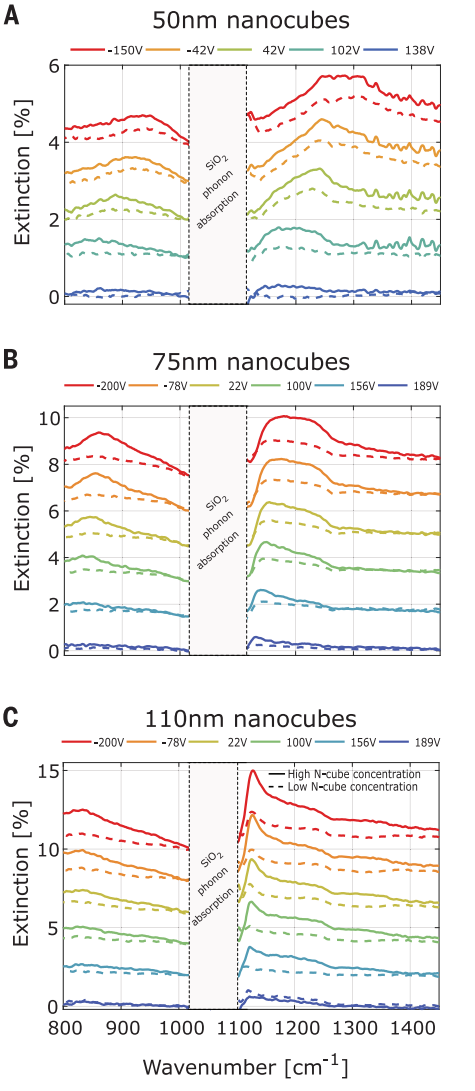
The effect of the polarization is shown for both polarized (solid curves) and unpolarized (dotted curves) illumination, and the two are nearly identical (Fig. 1C). Taking into account the fact that AGPs are transverse magnetic modes, this lack of preference on polarization in the response verifies the random nature of the nanocubes and implies a very weak interaction between neighboring nanocubes.

We further examined the device response by studying the gate-dependent extinction spectra obtained for different nanocube sizes and concentrations. The FTIR extinction spectrum was measured for three different nanocube sizes: 50 nm, 75 nm, and 110 nm (Fig. 2, A to C, respectively). For each nanocube size, two samples were fabricated and measured, one with a higher nanocube concentration (solid curves) and one with a lower nanocube concentration (dashed curves) (*21*). For all nanocube sizes, the resonances are seen to shift to higher frequencies with increasing graphene doping. An overall shift of the resonances to lower energies with increasing nanocube size can also be seen, which is attributed to the dispersive nature of AGPs (see below). In addition, a strong hybridization with the SiO_2 surface phonon occurs when the resonance frequency is close to that of the SiO_2 phonons; this is affected by the nanocube size and/or charge carrier density in graphene. This hybridization results in an enhancement of the peak around 1120 cm^{-1} as a result of the strong oscillator strength of the SiO_2 phonon.

The calculated and experimental dispersion relation (i.e., energy versus momentum dependence) shows good agreement (Fig. 2D). For simplicity, we calculate the dispersion for the layered structure without geometrical features (i.e., Ag/dielectric spacer/graphene/SiO_2; color map), as in (*16*), and the experimental momentum (colored stars) is calculated by $2\pi/L$, where L is the nanocube length. The resonances for the 75- and 110-nm nanocubes (green and red stars, respectively) lie very close to the SiO_2 phonons, thus exhibiting a more phonon-like nature and confirming the strong hybridization mentioned above. The resonances for the 50-nm nanocubes (black stars), however, are farther away in frequency from the SiO_2 phonons, thus displaying a more plasmon-like nature. This also explains the larger shift of the resonances with graphene doping that is obtained for the 50-nm nanocubes as compared to the 75- and 110-nm nanocubes, stemming from the fact that AGPs are doping-dependent whereas the SiO_2 phonons are not. It is also seen that the calculated dispersion lies close to the graphene Fermi velocity, $V_F = 1 \times 10^6$ m/s (dashed orange line), which corresponds to the lowest achievable velocity of AGPs and to the maximal wavelength confinement of $\lambda_0/300$. For the closest obtained experimental point (lower black star), we calculate an AGP velocity of ~1.42×10^6 m/s, denoting strong confinement of the optical field.

The results obtained in Fig. 2 further show that for different concentrations of the same nanocube size, one obtains the same spectral response but with different amplitude. On the other hand, different nanocube sizes generate an overall different spectral response. This implies that it is the properties of the single nanocube that determines the optical response of the device, whereas the amount (concentration) of nanocubes determines its amplitude. This is reinforced by the previous observation of the weak interaction between neighboring nanocubes. We can therefore conclude that a single nanocube/hBN/graphene structure actually acts as a single resonator. A further confirmation of the above conclusions is obtained from the 2D Fourier analysis of the scanning electron microscopy (SEM) images of the nanocubes, which distinctly show a lack of any other momentum contribution and/or periodical order in the samples' k-space, except for that of the single nanocube (*21*).

We note that in these samples, the measured fill factor, which is the percentage of area covered by the nanocubes, ranges from 3



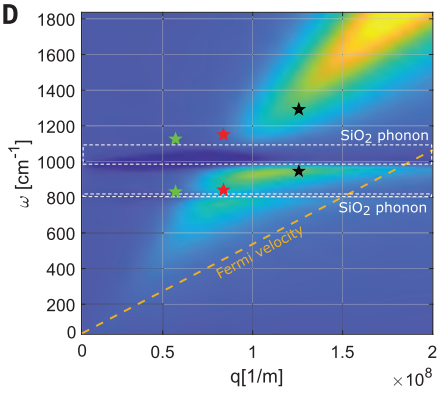


Fig. 2. Optical response for different nanocube sizes and concentrations. (**A** to **C**) Gate-dependent extinction for nanocube sizes of 50 nm (A), 75 nm (B), and 110 nm (C), showing the change in spectral response with nanocube dimension. The solid and dashed curves correspond to higher and lower nanocube concentrations, respectively, and show that the change in the nanocube concentration corresponds to a change in the amplitude of the response without affecting its spectrum. The data for different gate voltages have been shifted for clarity. For simplicity, no hBN capping layer is used in these samples. (**D**) Calculated dispersion relation for an Ag/2 nm dielectric spacer/graphene/SiO_2 structure, and extracted experimental results for nanocube sizes of 50 nm (black stars), 75 nm (red stars), and 110 nm (green stars). Graphene nonlocal conductivity was used, with Fermi level of 0.47 eV, lifetime of 10 fs, and T = 300 K for room temperature; the dashed orange curve represents the graphene Fermi velocity. The color scale corresponds to the imaginary part of the Fresnel reflection coefficient from the structure (the loss function).

to 13% (*21*). Yet even with such low fill factors, the optical response is still comparable in magnitude to that obtained for grating couplers, with an equivalent fill factor of 50 to 85% (*16*). Furthermore, this robust scheme does not require any lithographic processes or patterning, and thus it both preserves the graphene quality and removes any limitation on the device area. To increase the number of devices per sample, we limited the graphene area to stripes of 200 μm × 2 mm; however, centimeter-scale devices were also fabricated and measured (*21*).

To corroborate the single-resonator nature of the behavior, we performed 3D simulations of a single nanocube close to a graphene sheet, illuminated by a far-field free-space beam (*21*). The spatial distribution of the electric field $|E_y|$ shows that the field is mainly confined between the graphene and the nanocube (Fig. 3A), as expected for AGP modes (*15*, *16*, *20*). Figure 3, B and C, shows the obtained simulation results of the resonance frequency for different nanocube dimensions and Fermi levels, respectively. A linear dispersion relation is clearly seen in Fig. 3B, further corroborating the AGP nature of the resonance, together with a linear dependency on $\sqrt{E_F}$, which is characteristic of all plasmons in graphene. The field distribution obtained in Fig. 3A also exhibits the ability of the single structure to directly excite AGPs from the far field. However, there still remains an important question: What is the physical coupling mechanism that enables the far-field excitation of such high-momentum modes with a completely random structure that is built from single cavities?

It can be argued that the average inter-nanocube spacing might lead to a certain periodicity in the system, which can contribute momentum similar to a grating coupler. However, we see no change in the spectral response of different concentrations (Fig. 2), which is contradictory to what has been observed for periodical structures with different periods (*16*, *20*). This fact is further validated by the Fourier analysis of the nanocubes' distribution and by examining the response of grating couplers (*21*). We can therefore rule out the existence of any order-based momentum-matching condition. A further examination of the magnetic field distribution of the single structure reveals its coupling nature. Figure 3D shows the spatial distribution of the magnetic field for a single nanocube close to a graphene sheet, superimposed with the electric field lines. The lines around the graphene-nanocube interface form a loop that is correlated with a strong magnetic field in its center, which has the shape of a magnetic dipole resonance. In the radio-frequency regime, this type of behavior occurs when a rectangular metallic patch is placed above a grounded conductive plane, and is known as the patch antenna. It supports Fabry-Perot–like resonances and can be described by a magnetic surface current. For a conductive patch and ground plane, this electromagnetic response is not constrained to a specific spectral band. Indeed, metal nanocubes placed close to a gold surface, known as the nanocube-on-metal (NCoM) system, had been previously shown to act as magnetic resonators in the visible spectrum (*25*, *26*).

The optical response of the GPMR structure now becomes clear. Graphene, being a semimetal, can both act as a conductor and support AGPs when sufficiently doped. Thus, if a nanocube is placed in its vicinity, the illuminating far-field light can directly excite the GPMR patch antenna mode, which is associated with the excitation of an AGP between the nanocube and graphene, forming an optical cavity. The scalable nature of the patch antenna can also be observed in the GPMR spectral response for different nanocube sizes, where a larger nanocube size corresponds to a larger effective wavelength of the mode (see Fig. 2 and its discussion). We note that the remaining magnetic field within the top part of the nanocube in Fig. 3D is the part left unscreened by the graphene, owing to its small thickness and lower charge carrier density relative to a semi-infinite metal surface.

To corroborate the fact that the GPMR cavity is actually a graphene-based patch antenna in the MIR, we examined its scattering response. Owing to the similarity in the scattering response of patches and stripes (*25*, *27*), for simplicity of simulations we show in Fig. 4A the simulated 2D scattering of a single GPMR structure for several Fermi levels, ranging from 0.1 eV to the recently achieved 1.8 eV (*28*). Scattering resonances that are correlated with

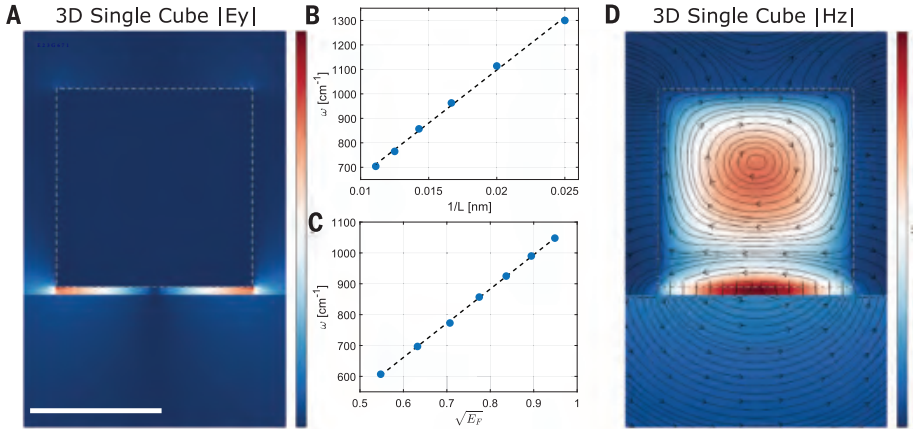


Fig. 3. 3D simulations and analysis of the generated electromagnetic fields of a single GPMR. (**A**) The simulated $|E_y|$ field distribution over the cross section of a single 3D nanocube. The intense electric field formed between the nanocube and the graphene is a signature of the generated AGPs (scale bar, 50 nm). (**B** and **C**) The expected linear dispersion of AGPs, as shown by the simulated resonance frequency obtained for different nanocube lengths L (40 to 90 nm) (B) and the square root of the Fermi level, $\sqrt{E_F}$ (0.3 to 0.9 eV) (C). (**D**) The simulated magnetic field distribution, $|H_z|$, superimposed with the electric field lines, showing the generation of a magnetic dipole resonance at the graphene/nanocube interface. Vacuum is used as the environment; spacer thickness, 3 nm.

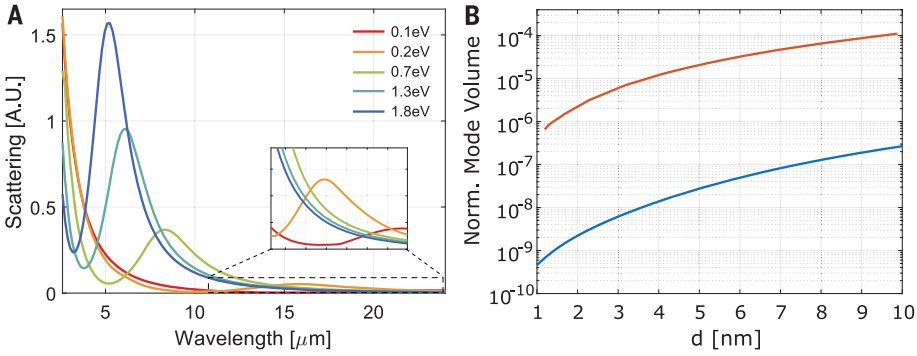


Fig. 4. GPMR antenna response and mode volume calculation. (**A**) Simulated 3D scattering from a single GPMR antenna for different graphene Fermi levels, showing that the scattering spectrum can be tuned from the far-infrared to almost the near-infrared. Inset: Zoom-in on the longer-wavelength spectrum. (**B**) Calculated normalized mode volume of the GPMR cavity in the MIR (blue curve) for different nanocube-graphene spacings d, and its comparison to the NCoM cavity in the visible spectrum (red curve), showing mode volumes that are smaller by four orders of magnitude with decreasing spacer thickness. See (*21*) for simulation details.

the AGP resonances can be observed, thus validating the GPMR antenna nature. The response can be tuned from the far-infrared almost to the near-infrared spectrum solely by changing the charge carrier density in the graphene. Such an ultrabroad spectral response is remarkable to obtain from a single antenna. Although a variety of graphene-based antennas have been proposed, mainly in the MIR and THz (*29*, *30*), they were found to be inefficient relative to their metallic counterparts. Thus, the GPMR might provide a different solution for realization of nanoantennas at these spectral bands.

Finally, we used the quasi-normal mode theory to study the mode volumes achieved by the GPMR cavity (*21*, *31*), neglecting nonlocal effects. Figure 4B shows the calculated normalized mode volume V_{GPMR}/λ_0^3 (where λ_0 is the free-space wavelength) of a single GPMR cavity with different hBN spacer thicknesses d (blue curve). Because similar plasmonic cavities in the visible spectrum have been shown to reach large confinement as well (*32*, *33*), the GPMR cavity mode volume is compared with the equivalent NCoM patch antenna in the visible spectrum, where graphene is replaced by a gold surface (*25*) (red curve). It can be seen that even though the GPMR wavelength is about an order of magnitude larger than the NCoM wavelength, it is able to achieve a normalized mode volume that is four orders of magnitude smaller, reaching a normalized mode volume of ~4.7×10^{10}. These remarkable values can be directly attributed to the unique properties of AGPs and their ability to confine light to very small dimensions.

Although we can experimentally thin down the hBN spacer until the monolayer case (*16*), we note that below thicknesses of 1 to 2 nm, strong nonlocal effects in the graphene should be introduced (*16*, *34*), given that we estimated the relative correction due to nonlocal effects to be ~23% in our experiment (*21*). In addition, we estimated the nonlocal response of the metal (*17*, *35*, *36*) and found it to be negligible above 1 nm (*21*). The inclusion of these requires a special treatment that cannot be introduced into our numerical simulations. Because it has been shown that AGPs can confine light to a monolayer hBN spacer and without increasing the losses (*16*), thinning down the spacer should continue and improve the achievable mode volume presented in Fig. 4B, accompanied by a spectral shift of the GPMR resonances and a saturation of the mode volume due to nonlocal effects (*16*).

The realization of single AGP cavities in the MIR spectrum with ultrasmall mode volumes, which are efficiently excited from the far field, provides a tunable platform for studying strong light-matter interactions in the MIR and THz spectra. One such important example is vibrational strong coupling (*37*) and its manipulation of chemical processes (*38*), which can be studied by introducing molecular substances or gas absorption layers between graphene and the nanocubes (*33*). In addition, it opens up possibilities for efficient AGP-based devices in this long-wavelength spectrum, such as photodetectors, biological and chemical sensing devices, and graphene-based tunable optical nanoantennas.

REFERENCES AND NOTES

1. B. Wunsch, T. Stauber, F. Sols, F. Guinea, *New J. Phys.* **8**, 318 (2006).
2. M. Jablan, H. Buljan, M. Soljačić, *Phys. Rev. B* **80**, 245435 (2009).
3. F. H. Koppens, D. E. Chang, F. J. García de Abajo, *Nano Lett.* **11**, 3370–3377 (2011).
4. A. N. Grigorenko, M. Polini, K. S. Novoselov, *Nat. Photonics* **6**, 749–758 (2012).
5. T. Low *et al.*, *Nat. Mater.* **16**, 182–194 (2017).
6. D. N. Basov, M. M. Fogler, F. J. García de Abajo, *Science* **354**, 195 (2017).
7. M. B. Lundeberg *et al.*, *Science* **357**, 187–191 (2017).
8. S. Chen *et al.*, *ACS Photonics* **4**, 3089–3097 (2017).
9. D. Rodrigo *et al.*, *Science* **349**, 165–168 (2015).
10. N. Rivera, I. Kaminer, B. Zhen, J. D. Joannopoulos, M. Soljačić, *Science* **353**, 263–269 (2016).
11. M. B. Lundeberg *et al.*, *Nat. Mater.* **16**, 204–207 (2017).
12. A. Woessner *et al.*, *Nat. Photonics* **11**, 421–424 (2017).
13. S. de Vega, F. J. García De Abajo, *ACS Photonics* **4**, 2367–2375 (2017).

14. K. A. Guerrero-Becerra, A. Tomadin, M. Polini, *Phys. Rev. B* **100**, 125434 (2019).
15. P. Alonso-González *et al*., *Nat. Nanotechnol.* **12**, 31–35 (2017).
16. D. Alcaraz Iranzo *et al*., *Science* **360**, 291–295 (2018).
17. E. J. Dias *et al*., *Phys. Rev. B* **97**, 245405 (2018).
18. J. B. Baxter, G. W. Guglietta, *Anal. Chem.* **83**, 4342–4368 (2011).
19. J. Haas, B. Mizaikoff, *Annu. Rev. Anal. Chem.* **9**, 45–68 (2016).
20. I.-H. Lee, D. Yoo, P. Avouris, T. Low, S.-H. Oh, *Nat. Nanotechnol.* **14**, 313–319 (2019).
21. See supplementary materials.
22. H. Yan *et al*., *Nat. Photonics* **7**, 394–399 (2013).
23. S. Dai *et al*., *Nat. Nanotechnol.* **10**, 682–686 (2015).
24. V. W. Brar *et al*., *Nano Lett.* **14**, 3876–3880 (2014).
25. A. Moreau *et al*., *Nature* **492**, 86–89 (2012).
26. J. J. Baumberg, J. Aizpurua, M. H. Mikkelsen, D. R. Smith, *Nat. Mater.* **18**, 668–678 (2019).
27. C. Wu, B. Neuner III, G. Shvets, J. John, A. Milder, B. Zollars, S. Savoy, *Phys. Rev. B* **84**, 075102 (2011).
28. K. Kanahashi *et al*., *npj 2D Mater. Appl.* **3**, 7 (2019).
29. J. M. Jornet, I. F. Akyildiz, *IEEE J. Sel. Areas Comm.* **31**, 685–694 (2013).
30. D. Correas-Serrano, J. S. Gomez-Diaz, arXiv 1704.00371 [cond-mat.mes-hall] (10 April 2017).
31. P. Lalanne, W. Yan, K. Vynck, C. Sauvan, J.-P. Hugonin, *Laser Photonics Rev.* **12**, 1700113 (2018).
32. G. M. Akselrod *et al*., *Nat. Photonics* **8**, 835–840 (2014).
33. R. Chikkaraddy *et al*., *Nature* **535**, 127–130 (2016).
34. B. A. Ferreira, B. Amorim, A. J. Chaves, N. M. R. Peres, *Phys. Rev. A* **101**, 033817 (2020).
35. M. Kamandar Dezfouli, C. Tserkezis, N. A. Mortensen, S. Hughes, *Optica* **4**, 1503 (2017).
36. Y. Yang *et al*., *Nature* **576**, 248–252 (2019).
37. M. Autore *et al*., *Light Sci. Appl.* **7**, 17172 (2018).
38. A. Thomas *et al*., *Science* **363**, 615–619 (2019).

ACKNOWLEDGMENTS

I.E. thanks E. J. C. Dias and F. Vialla for fruitful discussions. **Funding:** Supported by U.S. Air Force Office of Scientific Research (AFOSR) grants FA9550-12-1-0491 and FA9550-18-1-0187 (D.R.S.); AFOSR Foldable and Adaptive 2D Electronics Multidisciplinary University Research Initiative grant FA9550-15-1-0514 (J.-Y.H. and J.K.); the European Commission through the project "Graphene-Driven Revolutions in ICT and Beyond" (Ref. use CORE 3 reference, not CORE 2) (N.M.R.P.); COMPETE 2020, PORTUGAL 2020, FEDER, and the Portuguese Foundation for Science and Technology (FCT) through project POCI-01-0145-FEDER-028114 (N.M.R.P. and T.G.R.); and the Government of Catalonia through an SGR grant and the Spanish Ministry of Economy and Competitiveness through the "Severo Ochoa" Programme for Centres of Excellence in R and D (SEV-2015-0522), Fundacio Cellex Barcelona, Generalitat de Catalunya through the CERCA program, and Mineco grant Plan Nacional (FIS2016-81044-P), and Agency for Management of University and Research Grants (AGAUR) 2017 SGR 1656 (F.H.L.K.). Furthermore, the research leading to these results has received funding from the European Union Seventh Framework Programme under grant agreements 785219 and 881603 Graphene Flagship. This work was supported by the ERC TOPONANOP under grant agreement 726001. **Author contributions:** I.E. conceived the idea and performed simulations, experiments, and analysis of the results; D.A., A.K., and V.-V.P. fabricated devices and assisted in measurements; Z.H. and X.M.D. performed nanocube deposition and analysis; T.K. assisted in device fabrication; N.M.R.P., T.G.R., and J.-P.H. assisted in numerical simulations; J.-Y.H. and J.K. provided the monolayer hBN; and N.M.R.P., D.R.S., and F.H.L.K. supervised the project. All authors contributed to the writing of the manuscript. **Competing interests:** The authors have no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1219/suppl/DC1
Materials and Methods
Figs. S1 to S13
Table S1
References (*39*–*44*)

11 February 2020; accepted 24 April 2020
10.1126/science.abb1570

GEOPHYSICS

Sequencing seismograms: A panoptic view of scattering in the core-mantle boundary region

D. Kim[1]*, V. Lekić[1], B. Ménard[2], D. Baron[3], M. Taghizadeh-Popp[2]

Scattering of seismic waves can reveal subsurface structures but usually in a piecemeal way focused on specific target areas. We used a manifold learning algorithm called "the Sequencer" to simultaneously analyze thousands of seismograms of waves diffracting along the core-mantle boundary and obtain a panoptic view of scattering across the Pacific region. In nearly half of the diffracting waveforms, we detected seismic waves scattered by three-dimensional structures near the core-mantle boundary. The prevalence of these scattered arrivals shows that the region hosts pervasive lateral heterogeneity. Our analysis revealed loud signals due to a plume root beneath Hawaii and a previously unrecognized ultralow-velocity zone beneath the Marquesas Islands. These observations illustrate how approaches flexible enough to detect robust patterns with little to no user supervision can reveal distinctive insights into the deep Earth.

Seismic networks record millions of earthquake waveforms every year. The timing, amplitude, polarization, and frequency of their constituent arrivals contain precious information about seismic sources and deep Earth structures. For example, unexpected arrivals can reveal waves scattered or multipathed by heterogeneities within Earth, down to the core-mantle boundary (CMB). Mapping heterogeneities in the CMB region is important for understanding the fate of subducted slabs, the origin of hot spot volcanism, and the nature of primitive geochemical reservoirs (*1*). This mapping can be accomplished by identifying and interpreting scattered arrivals, which is challenging because it requires distinguishing seismic fluctuations from noise and contextualizing arrival amplitudes and timing in a regime where models are not currently available.

Traditionally, these challenges are overcome by focusing on a specific target area. This allows leveraging of geometric arrangements to identify robust signals and aid in their interpretation. A standard procedure is to arrange waveforms by epicentral distance or azimuth to reveal trends in arrivals not predicted by models of the interior. In Earth's CMB region, robustly identifying scattered waves has led to various discoveries, such as the D″ discontinuity (*2*), ultralow-velocity zones (ULVZs) (*3*), mega-ULVZs (*4*), and abrupt variations in wave speed across the boundaries of large low-shear-velocity provinces (LLSVPs) that imply compositional heterogeneity (*5*). In all cases, interpreting seismic waveforms would be difficult in isolation, without the context provided by seismic waves on nearby paths. Therefore, previous approaches have limited utility in poorly sampled regions and do not make use of the full statistical power of waveform datasets that span geographically diverse paths.

We conducted a large-scale, systematic search for seismic waves scattered by heterogeneity near the CMB across the Pacific basin. We focused on shear waves diffracting along the CMB (Sdiff), because they sample large areas at the base of the mantle. Waves scattered by heterogeneity in this region arrive after the main Sdiff phase, so the timing and amplitude of these Sdiff "postcursors" can constrain the location and nature of structures producing the scattering (*4*, *6*–*9*). Because energy reflected from the surface can complicate the identification of Sdiff postcursors, we restricted our attention to waveforms from deep earthquakes (*10*). This yielded a dataset of ~6000 transverse-component waveforms aligned on the main Sdiff arrival and deconvolved by synthetics computed for a one-dimensional (1D) preliminary reference Earth model (PREM) (*11*). When plotted by distance between the source and the receiver (Fig. 1A), postcursors cannot readily be identified. Ordering by azimuth only makes sense locally and is not useful when working with data spanning large geographic areas.

We made use of an unsupervised graph-based manifold learning algorithm, "the Sequencer" (*10*), that orders objects to minimize dissimilarities between neighbors as well as globally across the entire sequence. In our case, the objects are Sdiff waveforms and dissimilarity is given by the Wasserstein metric (also known as the earth mover's distance) between them (*10*). This approach can be used to reveal the main trend present in a dataset without requiring any model at all; it has been used in astronomy and has already led to the discovery of a trend relating the mass of supermassive

[1]Department of Geology, University of Maryland, College Park, MD 20742, USA. [2]Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA. [3]School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel.
***Corresponding author. Email: dk696@cornell.edu**

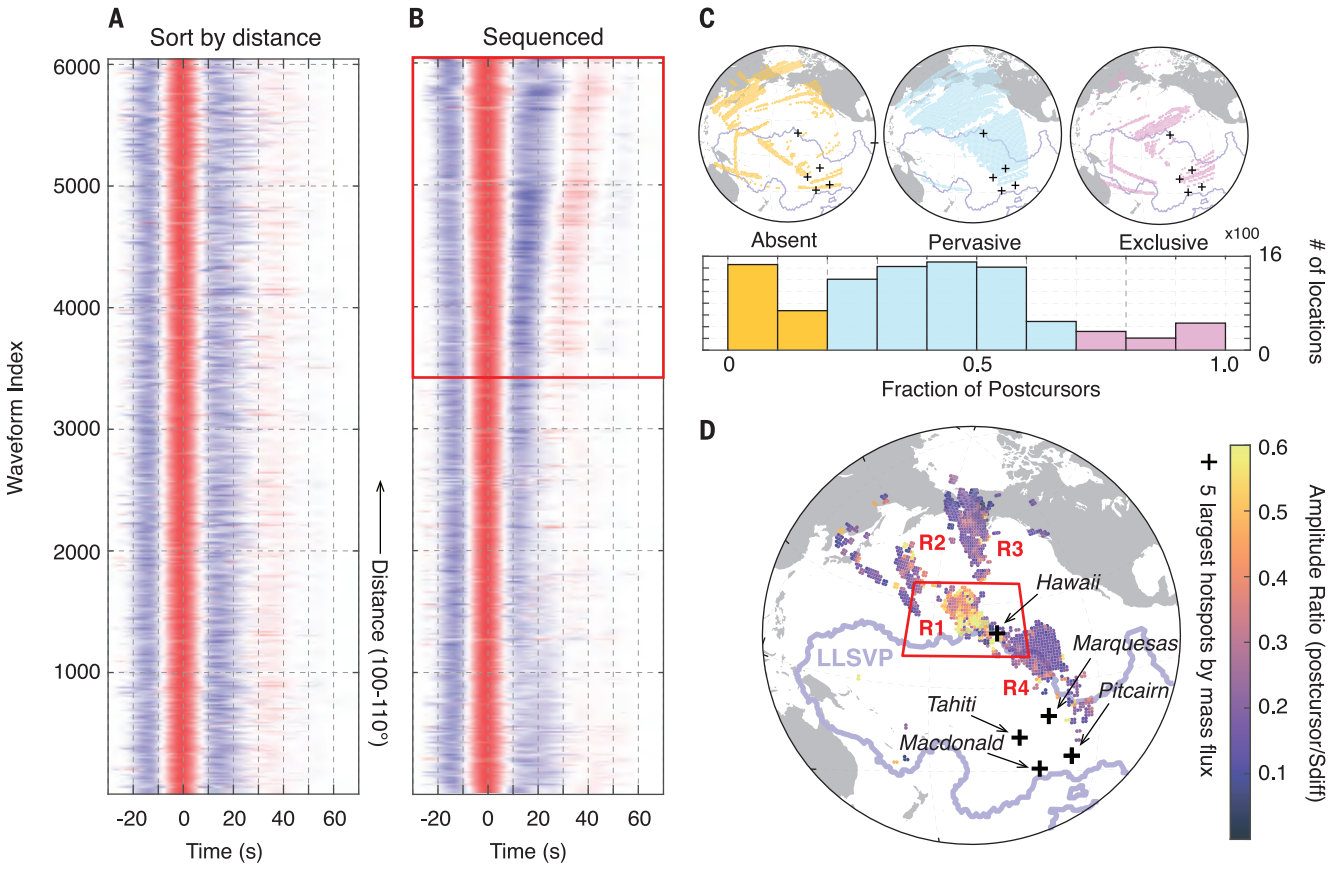


Fig. 1. Postcursors across the northern Pacific. A comparison of the source deconvolved Sdiff waveforms of moment magnitude (M_w) > 6.5 deep earthquakes (depth > 200 km) between 100° and 110° epicentral distance sorted by (**A**) distance and (**B**) the Sequencer. Sequencer ordering enables the identification of a substantial (~40% of all waveforms) subpopulation of Sdiff postcursors [red box in (B)]. (**C**) Histogram of fraction of waveforms with postcursors in 1° radius bins (bottom) and the geographic locations (top) where postcursors are absent (peach), pervasive (cyan), and exclusive (pink). (**D**) Stacks of postcursor amplitude relative to main Sdiff arrival, also averaged in 1° bins. The geographic extent of the Pacific LLSVP (*18*) is shown (light-blue contour), as are the five largest hot spots by mass flux (black crosses) (*35*). The region with the strongest postcursors is outlined in red. Regions labeled R1 to R4 are discussed in the main text.

black holes to the properties of their host galaxies (*12*). We used the Sequencer to optimally order our collection of Sdiff waveforms across the Pacific basin (Fig. 1B). We identified postcursors in >40% of waveforms, which indicates that postcursors are far more common than previously thought. Radial gradients in velocities cannot on their own explain the postcursors, regardless of ordering (figs. S1, B and D, and S2).

We then explored the geographic distribution of heterogeneities giving rise to the postcursors by binning the fraction of waveforms with postcursors in 1° radius bins (Fig. 1C). We found that heterogeneities large enough to scatter shear waves with periods >15 s are pervasive across the Pacific basin. In most locations, a substantial fraction (>0.3) of waves show postcursors (Fig. 1C, cyan). Postcursors are typically absent (Fig. 1C, peach) on paths that do not cross LLSVP boundaries and instead are confined either within the boundaries or well outside of them. Waveforms with and without postcursors seem to coexist at subwavelength scales (~160 km) in most locations (Fig. 1C). Many of the postcursors in the western and northern Pacific (Fig. 1D, R2 and R3) have large delay times (fig. S3B), indicating that they originate from distant scatterers and may travel through complex structures. These areas are suspected to host a partial melt created by paleo-slab from northwestern Pacific subduction zones (*13*) and a group of small ULVZ patches, along with slab debris beneath the northeastern boundary of the mid-Pacific LLSVP (*14*). Although large-amplitude postcursors appear to be associated with a few localities south of the Aleutians (Fig. 1D, R3), bootstrap error estimates show that these signals are not statistically significant (fig. S4). Notably, nearly all Sdiff wave sampling near Hawaii and the Marquesas Islands shows postcursors (Fig. 1C, pink). The amplitude of postcursors (with respect to the main Sdiff) in the Hawaiian region (Fig. 1D, R1) appears to be three times larger than the typical amplitude found throughout the Pacific basin (Fig. 1D, R2, R3, and R4).

When contextualized across all available data using the Sequencer, the region to the northwest of Hawaii stands out in terms of the prevalence of postcursors, their amplitude, and the spatial extent of the area associated with high-amplitude postcursors (Fig. 1, C and D). Therefore, we zoomed in on this region and performed a similar analysis of waveforms with turning points within 20° of Hawaii (Fig. 2A). The nature of such high-amplitude signals is discussed in the methods and materials section of the supplementary materials (*10*). By plotting the waveforms in the optimal order determined by the Sequencer, we readily identified a subpopulation featuring strong postcursors (Fig. 2B). Because of the limited geographic area, this population can also be identified by sorting the waveforms according to azimuth, with delay times increasing up to 40 s (fig. S5A). However, when visualized in the order determined by the Sequencer, the moveout of the postcursors is clarified, and coherent geographic patterns can be mapped.

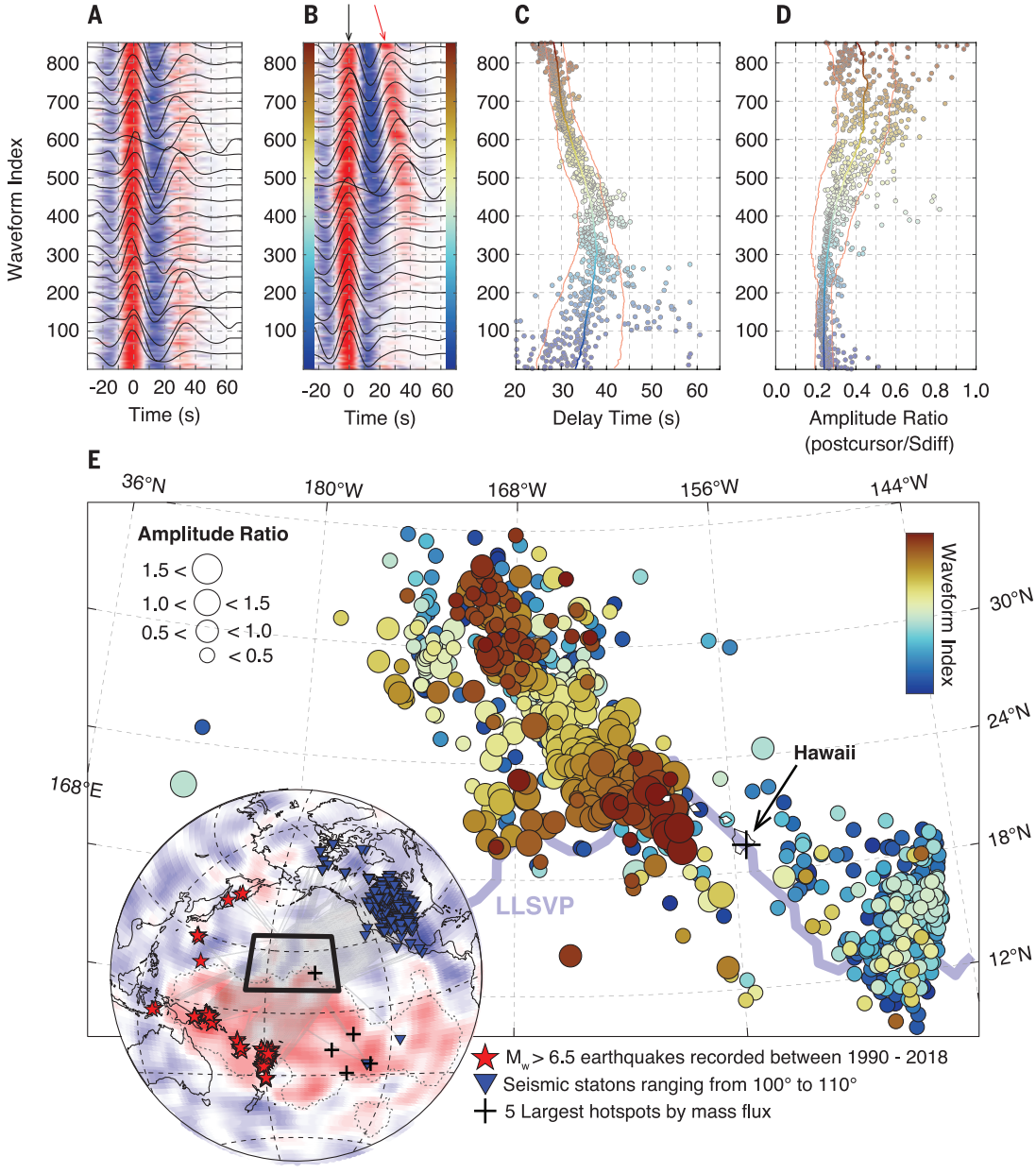


Fig. 2. Postcursors in the Hawaii region. Deconvolved Sdiff waveforms with turning points within 20° of Hawaii, sorted by (**A**) distance and (**B**) the Sequencer. Strong postcursors (red arrow) arrive after the main Sdiff phase (black arrow). Color bar represents sequence index and is used in later panels. The (**C**) delay time and (**D**) amplitude ratio of largest post-Sdiff arrivals, and (**E**) midpoints of their corresponding raypaths. The sequence reveals a coherent geographic pattern to the region northwest of Hawaii. Earthquakes and seismic stations are plotted as red stars and blue triangles (inset), respectively. The geographic extent of the Pacific LLSVP (*18*) is shown (purple contour), as are the five largest hot spots by mass flux (black crosses) (*35*).

By plotting each waveform's position in the sequence at the midpoint of its diffracting path, we found that waveforms with postcursors appear to the northwest of Hawaii, whereas those without appear predominantly to the southeast (Fig. 1D, R4). Moreover, the order identified by the Sequencer reveals a very distinctive spatial pattern (Fig. 2E), with waveforms appearing late in the sequence flanking a region in which midsequence waveforms cluster. This coherent pattern is reflected in both delay time and amplitude of the postcursors observed northwest of Hawaii (Fig. 2, C to E) and does not follow a simple azimuthal trend. Postcursor delay times gradually increase toward the center of the cluster, whereas large-amplitude postcursors (above ~0.5 with respect to the main Sdiff phase) are found mostly southward, near the northern edge of Pacific LLSVPs (Fig. 2E). The small number of weak postcursors detected to the southeast (Fig. 1D, R4) does not show coherent geographic trends, suggesting that they were not produced by a structure of the type located beneath Hawaii.

We also observed an anticorrelation between postcursor delay time and amplitude in the region (Fig. 2, C and D), consistent with expectations for a localized wave-speed anomaly. This observation motivated us to use the presence of such a correlation as a detector of localized structures. Thus, at each location across the Pacific basin, we estimated the slope and amplitude of the linear fit to the delay time and log amplitude of postcursors with diffracting paths located within 5° of the location. As expected, we found steep negative slopes in the vicinity of Hawaii, across a region that is substantially larger than anywhere else in the Pacific basin (Fig. 3A). In addition, we also detected a similar signature, with a slightly gentler slope, close to the Marquesas (Fig. 3A), indicating a previously unidentified localized wave-speed anomaly. Near both hot spots, the negative slopes are significant at a 99% level of confidence (Fig. 3B). Other locations with possible detections of localized wave-speed anomalies are in the vicinity of Alaska, Kamchatka, and along the northern edge of the Pacific LLSVP, but none are as clear as those near the Marquesas.

After using the Sequencer to reveal the leading trend in our waveform dataset and map the presence of heterogeneity in the Pacific region, we investigated the physical origin of the postcursor signals through waveform modeling. We carried out a systematic suite of wave-propagation simulations through candidate structures based on known features of the CMB region (*10*). Seismically imaged structures near the CMB span a wide range of sizes, from ~5000-km LLSVPs at one end (*15*, *16*) to ~100-km ULVZs on the other (*3*, *17*). So-called "mesoscale" structures have also been documented, including the Perm (*18*) and Kamchatka (*19*) anomalies, as well as unusually large ULVZs

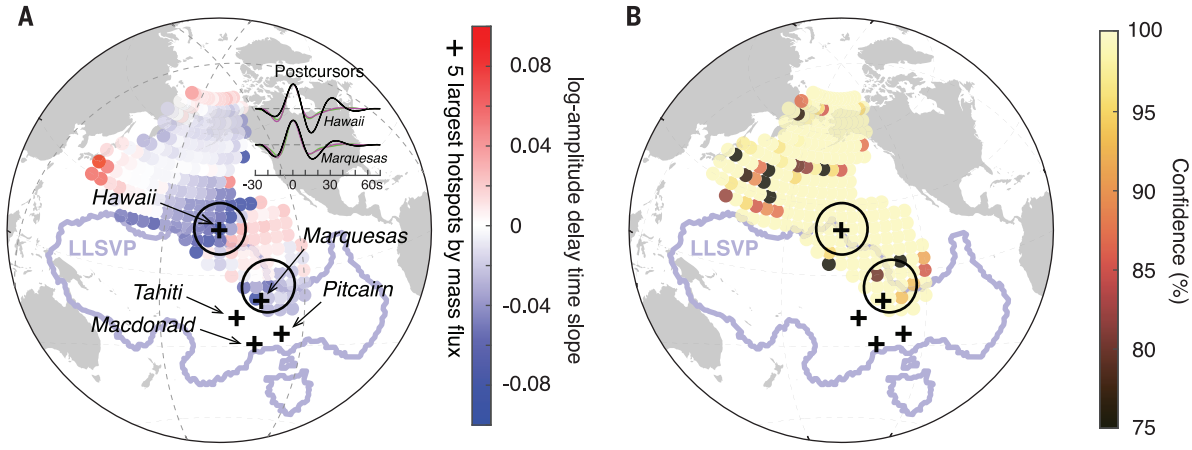


Fig. 3. Detection of localized structures. Map of (**A**) slope and (**B**) confidence range of the relationship between delay times and log amplitudes for postcursors identified in Fig. 1B and corresponding to raypaths that diffract within 5° of each location. Inset in (A) shows the average postcursor waveforms in Hawaii and the Marquesas (black), average of all nonpostcursor waveforms (green), and average of synthetic waveforms from PREM (magenta). Significance of these postcursor waveforms is discussed in the materials and methods (*10*) (fig. S10). Large negative value of the slope cluster around Hawaii and close to the Marquesas, indicating the presence of a localized source to the postcursors there. The geographic extent of the Pacific LLSVP (*18*) is shown (purple contour), as are the five largest hot spots by mass flux (black crosses) (*35*).

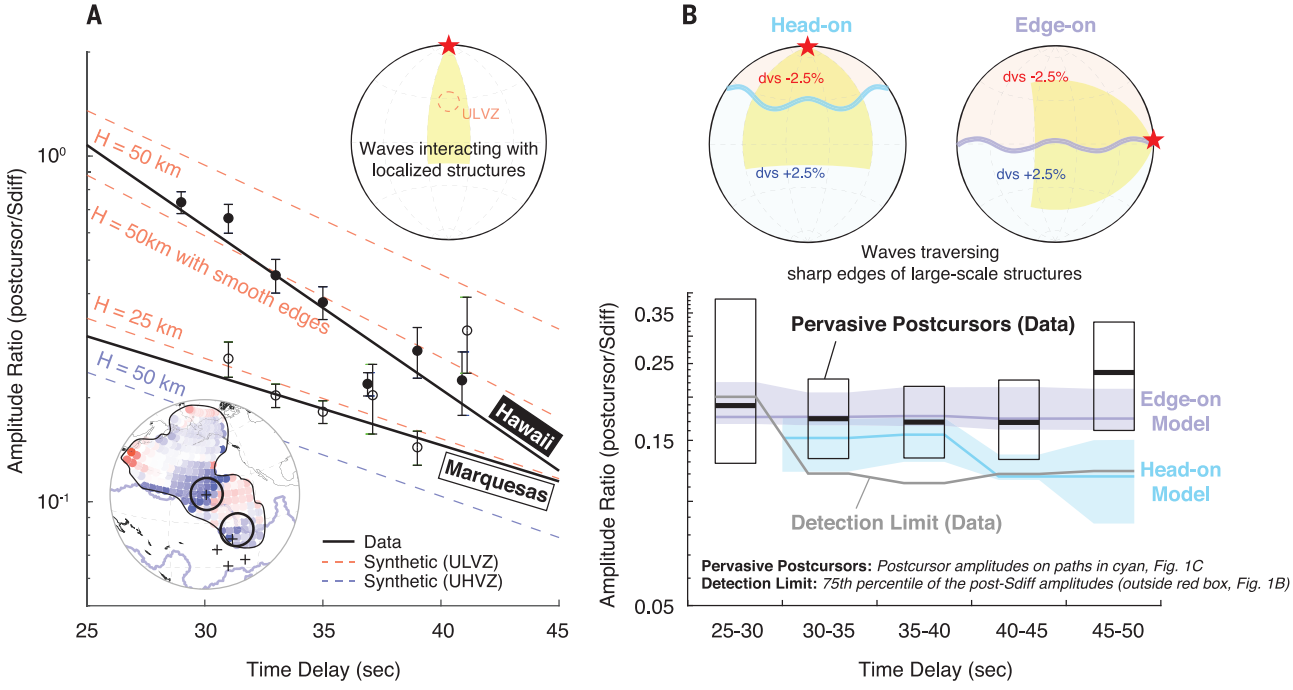


Fig. 4. Comparison of observed postcursor delay time versus log amplitude trends with predictions for candidate lowermost mantle structures. (**A**) Log amplitudes decrease strongly with delay time for postcursors near Hawaii (filled circles) and Marquesas (open circles) hot spots, in agreement with predictions (dashed lines) for three cylindrical ULVZ models (red). The prediction for an ultrahigh-velocity zone (UHVZ) model is plotted in blue. Upper inset shows illumination geometry of synthetic waveforms calculated for the cylindrical ULVZ models. Lower inset shows Fig. 3A and the diffraction coverage of the dataset (black outline). (**B**) For pervasive postcursors, log amplitude is observed to be constant with delay time (black boxplots) and agrees with the postcursors predicted for waves traversing a sharp LLSVP boundary (5% *Vs* contrast) edge-on (lavender). Head-on geometries (light blue) underpredict the amplitudes. Detection limit is defined by the 75th percentile of the post-Sdiff amplitude in nonpostcursor waveforms (outside the red box in Fig. 1B).

(*4*, *6*, *20*). Therefore, we explored two types of candidate structures: (i) cylindrical regions of reduced or elevated shear-wave velocity (V_S) with dimensions reminiscent of ULVZs and (ii) undulating boundaries of large low-V_S regions reminiscent of LLSVPs (Fig. 3). We explored the effects of the lateral abruptness of velocity changes across the boundaries of both types of structures.

Our waveform simulations confirmed that postcursor log amplitude decays linearly with delay time up to 20° away from a low-velocity cylinder (Fig. 4A, inset). The slope of this decay (Fig. 3) is controlled by geometric spreading and seismic attenuation, while cylinder height and width and the magnitude of its velocity reduction determine the amplitude. A high-velocity cylinder produces postcursors that are four to five times weaker, with similar log amplitude decay (Fig. 4A). Differences in effects of fast and slow anomalies are discussed in the materials and methods (*10*). Trade-offs among these physical parameters make it impossible to uniquely map slope and zero-delay amplitude to anomaly size, shape, and V_S reduction (Fig. 4A and fig. S6). Nevertheless, we found that a published model for the Hawaii ULVZ (*4*, *7*, *8*) (fig. S7) with a 20% reduction in V_S [height (H) = 25 km] (Fig. 4A, dashed red lines) matched the amplitude and delay time of the Marquesas postcursors. This indicates the presence of a mega-ULVZ beneath the Marquesas, although our dataset does not constrain its precise location and characteristics.

However, this model fails to reproduce the large postcursor amplitudes we observe near Hawaii. Instead, a narrower but taller (H = 600 km) low-velocity body representing a deeply rooted plume conduit can match the amplitude and delay time of Hawaii postcursors, as does a 50-km-tall cylinder with more gradual boundaries (H = 50 km, with smooth edges) (Fig. 4A). Waveforms computed in the plume root model for an earthquake which samples the Hawaii region agree well with observations (M_B in figs. S8 and S9). Such a narrow (<500-km-wide) mantle plume is not resolvable by travel-time tomography (*21*) but should become visible in full waveform inversions (*22*) incorporating the postcursor waveforms. Although more complicated models cannot be ruled out, such as a 100-km-tall extremely low V_S (25% reduction) structure embedded in the northern edge of Pacific LLSVPs (*7*), we prefer the plume root structure to explain the amplitudes and waveforms of Hawaii postcursors.

Away from Hawaii and the Marquesas, postcursor amplitude (Fig. 4B, black) is constant with delay time and significantly greater (fig. S10) than the largest post-Sdiff signal in waveforms that do not show a postcursor when sequenced (Fig. 4B, gray). Postcursor amplitudes cannot be attributed to deconvolution artifacts (*10*) (fig. S10). The spatial pattern of postcursor amplitude variations has a characteristic length scale of ~3000 km (fig. S11), which is similar to the Fresnel width for a scattered arrival with a delay time of 30 to 50 s. We ruled out that the structure beneath Hawaii alone produces the postcursors observed across the northern Pacific, because our wavefield simulations indicate that regardless of height, this structure cannot produce postcursors of sufficiently large amplitude on paths >20° away from its center (fig. S8). Rather, two scenarios can plausibly explain postcursors identified by the Sequencer across the northern Pacific: (i) scattering from multiple smaller anomalies distributed across the region and (ii) scattering or multipathing across sufficiently laterally abrupt boundaries of the LLSVPs.

Multiple anomalies distributed across the Pacific could have complex geometries or comprise multiple ULVZs with various geometries and sizes (*23*). A conglomerate of individual ULVZs that are smaller than our wavelength and ubiquitous across the region may produce the pervasive postcursors we observed. Whether these anomalies involve partial melting (*24*) or compositionally distinct, low V_S materials (*25*) should be further examined using various body-wave phases at higher frequencies (*26*).

Alternatively, sharp edges and complexities associated with the LLSVP (*27*, *28*) can broaden shear-wave pulses (Fig. 1B) that propagate nearby (*29*) or result in multipathing (*9*) and scattering that can produce postcursors. As demonstrated by our wavefield simulations, the postcursors resulting from waves approaching the edges of such a large-scale structure head-on (Fig. 4B, light blue) generally produce weaker-than-observed postcursors. On the other hand, postcursors generated from the waves that transmit obliquely across the same 5% impedance contrast (Fig. 4B, lavender) better reproduce observed amplitudes. The effects of the smoothness of those edges (i.e., sharpness of the velocity along the boundary) were negligible. To test postcursor generation with a more realistic model, we have sharpened the bottom 600 km of the SEMUCB-WM1 tomographic model (*30*) to introduce sharper edges for the LLSVPs (M_C in fig. S8B). Synthetic waveforms produced by this sharpened model provide a moderately improved fit to the postcursors observed outside the Hawaii region (fig. S9) but could not generate the high-amplitude postcursors near Hawaii (Fig. 4A and fig. S8).

Exploring a large dataset with the Sequencer enables a data-driven analysis of seismic waveforms without any prior expectations. We anticipate this approach to be useful for many types of datasets beyond seismograms. Often, observed phenomena are driven by a leading effect or parameter. In such cases, there should exist a 1D manifold representing this underlying trend, even if it exhibits a complex behavior. Manifold learning techniques such as the Sequencer can reveal this leading trend from complex data, which is especially useful in the absence of theoretical guidance. When the leading trend is already known, it can be removed before sequencing, as we did by deconvolving the 1D Earth response from the seismograms. In our case, the trends identified by the Sequencer end up being surprisingly simple. They could also be obtained by plotting the delay time and amplitude of a Gaussian fit to the postcursors. However, this simplicity was not obvious before "sequencing" the dataset. In other words, once one knows what to look for, revealing the trends presented in this work is straightforward and does not require the Sequencer algorithm.

Incorporating higher-frequency waveforms would enable us to constrain structures of even smaller scale. These waveforms are harder to interpret but should not pose a challenge to our manifold learning approach. Our wavefield simulations indicate that, given their frequency content (15 to 100 s), our data cannot resolve velocity variations at scales smaller than ~50 km. Although geometry of seismic illumination to the lowermost mantle will remain more or less the same, Sequencer-based approaches for identifying trends and anomalous signals in portions of seismograms that host various seismic phases will advance our understanding of deep Earth structures.

Our systematic postcursor search in the Pacific basin, interpreted with insights from waveform simulations, reveals that heterogeneity capable of producing postcursors is widespread in the CMB region. We found that strong postcursors exhibit an anticorrelation between delay time and log amplitude that can be detected locally. The two strongest postcursor signals in the Pacific are related to the Marquesas and a broad region near Hawaii. The previously unknown localized anomaly near the Marquesas has dimensions similar to that of a mega-ULVZ (*4*). The anomaly beneath Hawaii is unique in the Pacific basin, producing postcursors that can only be explained by a structure either substantially larger than a mega-ULVZ or by a plume conduit. Weaker postcursors are observed throughout the Pacific basin and do not exhibit a correlation between delay time and amplitude; they are best explained by scattering from laterally abrupt edges of LLSVPs (5% V_S contrast). The origins of and relationships among these CMB region structures remain controversial (*31*). Nevertheless, recent analyses suggest that LLSVPs may host relatively undegassed geochemical reservoirs (*32*). Mega-ULVZs, on the other hand, must involve either exotic compositions or reflect the presence of partial melt (*24*, *33*), and they have been proposed to

host primitive geochemical signatures predating the moon-forming impact (*34*). Therefore, our discovery of a mega-ULVZ beneath the Marquesas offers a test of this hypothesis.

REFERENCES AND NOTES

1. A. M. Dziewonski, B. A. Romanowicz, *Treatise on Geophysics* **1**, 1–28 (2015).
2. T. Lay, Q. Williams, E. J. Garnero, *Nature* **392**, 461–468 (1998).
3. E. J. Garnero, S. P. Grand, D. V. Helmberger, *Geophys. Res. Lett.* **20**, 1843–1846 (1993).
4. S. Cottaar, B. Romanowicz, *Earth Planet. Sci. Lett.* **355–356**, 213–222 (2012).
5. L. Wen, P. Silver, D. James, R. Kuehnel, *Earth Planet. Sci. Lett.* **189**, 141–153 (2001).
6. K. Yuan, B. Romanowicz, *Science* **357**, 393–397 (2017).
7. A. To, Y. Fukao, S. Tsuboi, *Phys. Earth Planet. Inter.* **184**, 119–133 (2011).
8. A. To, Y. Capdeville, B. Romanowicz, *Phys. Earth Planet. Inter.* **256**, 26–36 (2016).
9. A. To, B. Romanowicz, Y. Capdeville, N. Takeuchi, *Earth Planet. Sci. Lett.* **233**, 137–153 (2005).
10. Materials and methods are available as supplementary materials.
11. A. M. Dziewonski, D. L. Anderson, *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).
12. D. Baron, B. Ménard, *Mon. Not. R. Astron. Soc.* **487**, 3404–3418 (2019).
13. Y. Xu, K. D. Koper, *Geophys. Res. Lett.* **36**, L17301 (2009).
14. D. Sun *et al.*, *Geophys. Res. Lett.* **46**, 3142–3152 (2019).
15. S. Cottaar, V. Lekic, *Geophys. J. Int.* **207**, 1122–1136 (2016).
16. E. J. Garnero, A. K. McNamara, S. H. Shim, *Nat. Geosci.* **9**, 481–489 (2016).
17. S. Yu, E. J. Garnero, *Geochem. Geophys. Geosyst.* **19**, 396–414 (2018).
18. V. Lekic, S. Cottaar, A. Dziewonski, B. Romanowicz, *Earth Planet. Sci. Lett.* **357–358**, 68–77 (2012).
19. Y. He, L. Wen, T. Zheng, *Geophys. Res. Lett.* **41**, 7073–7079 (2014).
20. M. S. Thorne, E. J. Garnero, G. Jahnke, H. Igel, A. K. McNamara, *Earth Planet. Sci. Lett.* **364**, 59–67 (2013).
21. R. Maguire, J. Ritsema, M. Bonnin, P. E. van Keken, S. Goes, *J. Geophys. Res. Solid Earth* **123**, 384–400 (2018).
22. S. W. French, B. Romanowicz, *Nature* **525**, 95–99 (2015).
23. C. Zhao, E. J. Garnero, M. Li, A. McNamara, S. Yu, *J. Geophys. Res. Solid Earth* **122**, 1198–1220 (2017).
24. Q. Williams, E. J. Garnero, *Science* **273**, 1528–1530 (1996).
25. W. L. Mao *et al.*, *Science* **312**, 564–565 (2006).
26. T. Lay, *Treatise on Geophysics* **1**, 619–654 (2015).
27. S. Ni, E. Tan, M. Gurnis, D. Helmberger, *Science* **296**, 1850–1852 (2002).
28. T. Lay, J. Hernlund, E. J. Garnero, M. S. Thorne, *Science* **314**, 1272–1276 (2006).
29. C. Zhao, E. J. Garnero, A. K. McNamara, N. Schmerr, R. W. Carlson, *Earth Planet. Sci. Lett.* **426**, 143–153 (2015).
30. S. W. French, B. A. Romanowicz, *Geophys. J. Int.* **199**, 1303–1327 (2014).
31. A. K. McNamara, *Tectonophysics* **760**, 199–220 (2019).
32. C. D. Williams, S. Mukhopadhyay, M. L. Rudolph, B. Romanowicz, *Geochem. Geophys. Geosyst.* **20**, 4130–4145 (2019).
33. J. K. Wicks, J. M. Jackson, W. Sturhahn, *Geophys. Res. Lett.* **37**, L15304 (2010).
34. A. Mundl *et al.*, *Science* **356**, 66–69 (2017).
35. B. Steinberger, *J. Geophys. Res. Solid Earth* **105**, 11127–11152 (2000).

ACKNOWLEDGMENTS

We thank B. Romanowicz, A. To, S. Cottaar, N. Schmerr, and N. Guttenberg for insightful discussions that helped improve this manuscript. We acknowledge the thorough and thoughtful reviews from three anonymous reviewers, which also greatly improved the manuscript. **Funding:** This work is supported by Packard Foundation Fellowships to V.L. and B.M. and NSF-EAR1352214 grant to V.L. Development of the Sequencer was supported by the generosity of Eric and Wendy Schmidt by recommendation of the Schmidt Futures program. **Authors contributions:** D.K., V.L., and B.M. provided the intellectual framework, designed the research, and wrote the manuscript. D.K. performed the data analysis and computations and prepared the figures and technical details. B.M. and D.B. developed the Sequencer, while M.T.-P. implemented the online Sequencer used for this research. All authors helped edit the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Seismic data used in this manuscript are available through the IRIS Data Management Center (DMC). The source code of the Sequencer, together with implementation details, can be found at https://github.com/dalya/Sequencer/. An online version of the algorithm is available at http://sequencer.org.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1223/suppl/DC1
Materials and Methods
Figs. S1 to S16
References (*36–44*)

14 January 2020; accepted 4 May 2020
10.1126/science.aba8972

ELECTROCHEMISTRY

Chloride-mediated selective electrosynthesis of ethylene and propylene oxides at high current density

Wan Ru Leow[1]*, Yanwei Lum[1,2]*, Adnan Ozden[3], Yuhang Wang[1], Dae-Hyun Nam[1], Bin Chen[1], Joshua Wicks[1], Tao-Tao Zhuang[1], Fengwang Li[1], David Sinton[3], Edward H. Sargent[1]†

Chemicals manufacturing consumes large amounts of energy and is responsible for a substantial portion of global carbon emissions. Electrochemical systems that produce the desired compounds by using renewable electricity offer a route to lower carbon emissions in the chemicals sector. Ethylene oxide is among the world's most abundantly produced commodity chemicals because of its importance in the plastics industry, notably for manufacturing polyesters and polyethylene terephthalates. We applied an extended heterogeneous:homogeneous interface, using chloride as a redox mediator at the anode, to facilitate the selective partial oxidation of ethylene to ethylene oxide. We achieved current densities of 1 ampere per square centimeter, Faradaic efficiencies of ~70%, and product specificities of ~97%. When run at 300 milliamperes per square centimeter for 100 hours, the system maintained a 71(±1)% Faradaic efficiency throughout.

In the United States, chemical manufacture accounts for 28% of total industrial energy demand (*1*). At present, this demand is largely met by the consumption of fossil fuels, resulting in substantial carbon dioxide (CO_2) emissions (*2*, *3*); a recent report showed that the plastics industry alone releases 1.8 billion metric tons of CO_2 per year and that replacing fossil fuels–based production methods with ones powered with renewable energy offers a route to reduce net greenhouse gas emissions associated with plastics manufacture (*4*).

One attractive strategy involves the development of electrochemical systems that produce the necessary raw materials by using renewable electricity (*5–8*). Ethylene oxide is used in the manufacture of plastics, detergents, thickeners, and solvents (*9*) and is among the world's top 15 most abundantly produced chemicals at ~20 million metric tons per year (*10*, *11*). At present, it is manufactured through the silver (Ag)–catalyzed direct oxidation of ethylene at high temperature and pressure (200° to 300°C and 1 to 3 MPa). This process generates 0.9 tons of CO_2 per ton of ethylene oxide produced, with more than half attributed to the complete combustion of ethylene and the balance arising from temperature-control units (Fig. 1A), which today are fossil fuel–powered (*12*).

The contribution from the complete oxidation of ethylene all the way to CO_2 is a result of the mechanism relied on in direct ethylene oxidation: One ethylene reacts with one oxygen (O) atom from Ag-adsorbed dioxygen to form ethylene oxide, leaving behind the remaining Ag-adsorbed O atom, and this oxidizes ethylene all the way to CO_2. It takes six O atoms for the complete combustion of one ethylene molecule; this means that, in the best case, for every six molecules of ethylene converted to ethylene oxide, one molecule of ethylene will be completely combusted to CO_2, thus limiting the product specificity to a theoretical upper limit of 85.7% (*9*). Here, specificity refers to the percentage of reacted substrate (ethylene) that goes toward the desired product (ethylene oxide).

Additionally, fossil-powered systems are typically used to maintain a stable temperature profile in order to suppress thermal runaway in this highly exothermic reaction.

We pursued an electrochemical approach to the production of ethylene oxide, to address

[1]Department of Electrical and Computer Engineering, University of Toronto, 35 St. George Street, Toronto, ON M5S 1A4, Canada. [2]Institute of Materials Research and Engineering, Agency for Science, Technology and Research (A*STAR), 2 Fusionopolis Way, Innovis, Singapore 138634, Singapore. [3]Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, ON M5S 3G8, Canada.
*These authors contributed equally to this work.
†**Corresponding author. Email: ted.sargent@utoronto.ca**

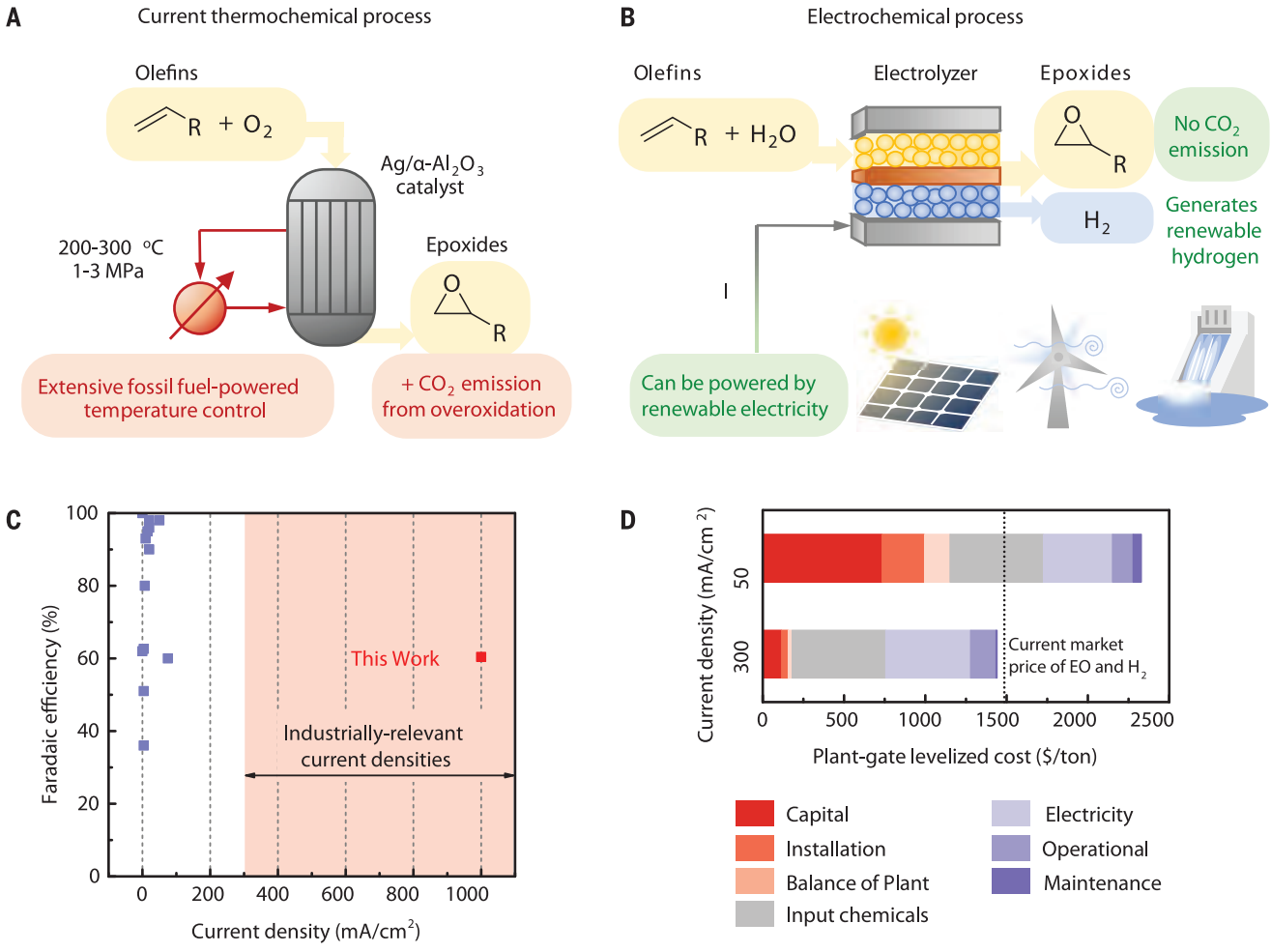


Fig. 1. Electrosynthesis of ethylene oxide by using renewable energy. (**A** and **B**) Schematics illustrating (A) the industrial thermochemical system and (B) the proposed electrochemical system. (**C**) Reported current densities and Faradaic efficiencies for other anodic partial oxidation reactions in the literature (blue squares) (*15*–*24*, *34*–*36*). Data for the system demonstrated in this work are shown for comparison (red square). (**D**) Breakdown of costs at current densities of 50 and 300 mA/cm^2, as calculated from a TEA. TEA calculation details are provided in the supplementary text and fig. S1.

both the first problem (limited product specificity) and the second (the need to reduce the use of fossil energy sources in powering systems). We sought a highly selective electrochemical route, pursuing the electrooxidation of ethylene to ethylene oxide with high product specificity and Faradaic efficiency under ambient conditions as a means to contribute to lowering CO_2 emissions in the production of this chemical (Fig. 1B) (*13*, *14*).

The electrosynthesis of ethylene oxide involves the partial oxidation of ethylene, an anodic reaction. Reactions of this nature at high current density and Faradaic efficiency are hampered by two challenges. First, the large positive potentials necessary can lead to uncontrolled over-oxidation, generating undesired by-products such as CO_2. Currently, reported anodic upgrading reactions such as the oxidation of 5-hydroxymethylfurfural (*15*–*17*), alcohol (*18*–*20*), and glycerol (*21*–*23*) are conducted at low current densities (<100 mA/cm^2) to maximize Faradaic efficiencies toward the target product (Fig. 1C). However, the production of industrially relevant quantities of the product at such low current densities would require unreasonably high electrolyzer surface areas, leading to high capital costs per unit of productivity (Fig. 1D). Second, if the reactant (ethylene) has limited solubility in the aqueous electrolyte, the system quickly becomes mass transport–limited, resulting in poor Faradaic efficiency at high current densities.

We took the view that a selective production strategy could avoid directly oxidizing the organic reactant molecules on the electrode surface so as to prevent over-oxidation at high current densities. We reasoned that a redox mediator that facilitates the indirect exchange of electrons between the electrode and the substrate molecules would allow this. Furthermore, in such a scheme, the space in which the reaction takes place is not limited to the planar electrode:electrolyte interface but extends into the bulk electrolyte, constituting an extended heterogeneous:homogeneous interface (Fig. 2A) that overcomes mass transport limitations. Using this strategy, we demonstrated ethylene oxide production at high current density (up to 1 A/cm^2), Faradaic efficiency (~70%), and product specificity (~97%).

Initially, we attempted to oxidize ethylene directly to ethylene oxide using a nanostructured palladium (Pd) anode (fig. S2A). This approach was based on a recent study in which olefins such as propylene were oxidized at low current densities (*24*) but did not translate to the high current densities; at 300 mA/cm^2, a negligible Faradaic efficiency was obtained toward ethylene oxide (fig. S2B). Operating at this high current density resulted in dissolution of the Pd anode, as can be observed from the rapidly increasing potential with time (fig. S2C).

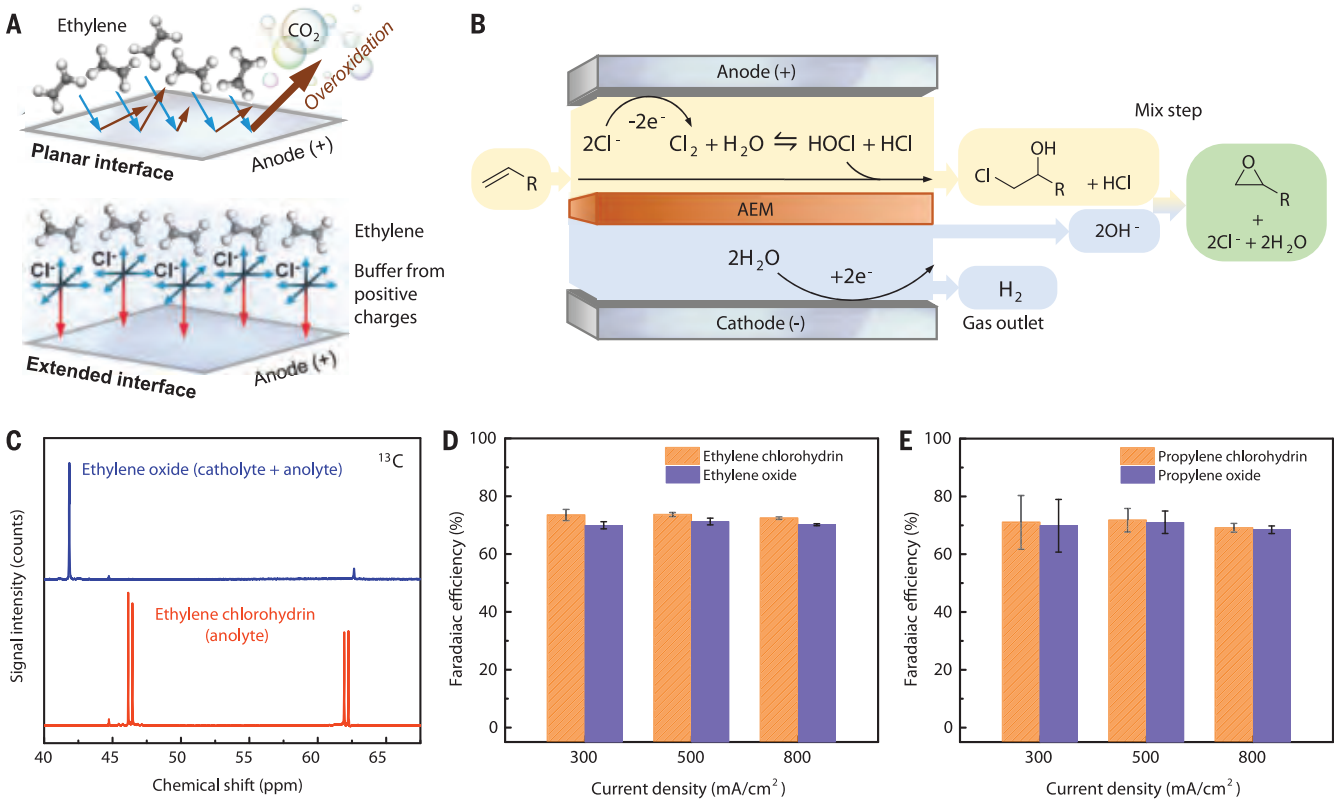


Fig. 2. Selective ethylene oxide production from ethylene enabled by an extended heterogeneous:homogenous interface. (**A**) Schematic illustrating ethylene oxidation at planar versus extended interfaces. (**B**) Schematic of the ethylene-to–ethylene oxide electrochemical system. A detailed schematic of the electrolyzer is available in fig. S3. (**C**) ^{13}C NMR spectra of ethylene oxide and ethylene chlorohydrin. (**D**) Faradaic efficiencies of ethylene oxide and ethylene chlorohydrin at different current densities. (**E**) Faradaic efficiencies of propylene oxide and propylene chlorohydrin at different current densities. The error bars correspond to the standard deviation of three independent measurements.

Additionally, the use of organic mediators such as TEMPO [(2,2,6,6-tetramethylpiperidin-1-yl)oxyl] (*15*, *25*) and NHPI (*N*-hydroxyphthalimide) (*26*, *27*)—a method to obtain high selectivities for partial oxidation products at the anode—failed in the generation of ethylene oxide and yielded instead only small amounts of acetate (Fig. 2B).

Historically, ethylene oxide was produced from ethylene through the addition of aqueous chlorine to form ethylene chlorohydrin, followed by dehydrochlorination with calcium hydroxide to generate ethylene oxide (*9*). However, the high cost associated with consuming stoichiometric amounts of chlorine and hydroxide, as well as disposal of the water medium, has lowered interest in this process. We postulated that chloride (Cl^-) could be a redox mediator at the anode with extended heterogeneous:homogeneous interfaces. The Cl^- could thereby buffer ethylene from uncontrolled oxidation and facilitate ethylene oxide production.

This idea was tested in a flow-cell setup with 1.0 M potassium chloride (KCl) electrolyte, in which ethylene was continuously sparged into the anolyte, with platinum (Pt) foil as the working electrode (anode), nickel (Ni) foam as the counter electrode (cathode), and Ag/AgCl (3.0 M KCl) as the reference electrode (fig. S3). The geometric surface area of the anode used was 1 cm^2, and the catholyte and anolyte volumes were both 25 ml. An anion exchange membrane (AEM) separated the anolyte and catholyte chambers. To determine the Faradaic efficiency, product quantification was carried out by means of high-performance liquid chromatography (HPLC) (supplementary materials, materials and methods). Unless otherwise stated, all electrolysis experiments were run for a duration of 1 hour.

Experiments were first carried out at 300 mA/cm^2. On the basis of prior studies (*28*), in this process Cl^- is oxidized to Cl_2 at the Pt anode (Eq. 1; *e* is the charge on the electron), which disproportionates in the aqueous environment to form hypochlorous and hydrochloric acid (HOCl and HCl, respectively) (Eq. 2) (*29*, *30*); HOCl then reacts with ethylene dissolved in the electrolyte to form ethylene chlorohydrin ($HOCH_2CH_2Cl$) (Eq. 3). Because HCl is not consumed, the pH of the anolyte becomes acidic over the course of the electrolysis (pH 1.1)

$$\text{Anode: } 2Cl^- \rightarrow Cl_2 + 2e^- \quad (1)$$

$$Cl_2 + H_2O \leftrightharpoons HOCl + HCl \quad (2)$$

$$C_2H_4 + HOCl \rightarrow HOCH_2CH_2Cl \quad (3)$$

$$\text{Cathode: } 2H_2O + 2e^- \rightarrow H_2 + 2OH^- \quad (4)$$

$$\text{Mixing step: } HOCH_2CH_2Cl + OH^- \rightarrow C_2H_4O + Cl^- + H_2O \quad (5)$$

$$HCl + OH^- \rightarrow Cl^- + H_2O \quad (6)$$

$$\text{Overall: } C_2H_4 + H_2O \rightarrow C_2H_4O + H_2 \quad (7)$$

The final step involves addition of hydroxide (OH^-), which then reacts with ethylene chlorohydrin to yield the desired ethylene oxide and regenerate Cl^- (*28*). The hydrogen evolution

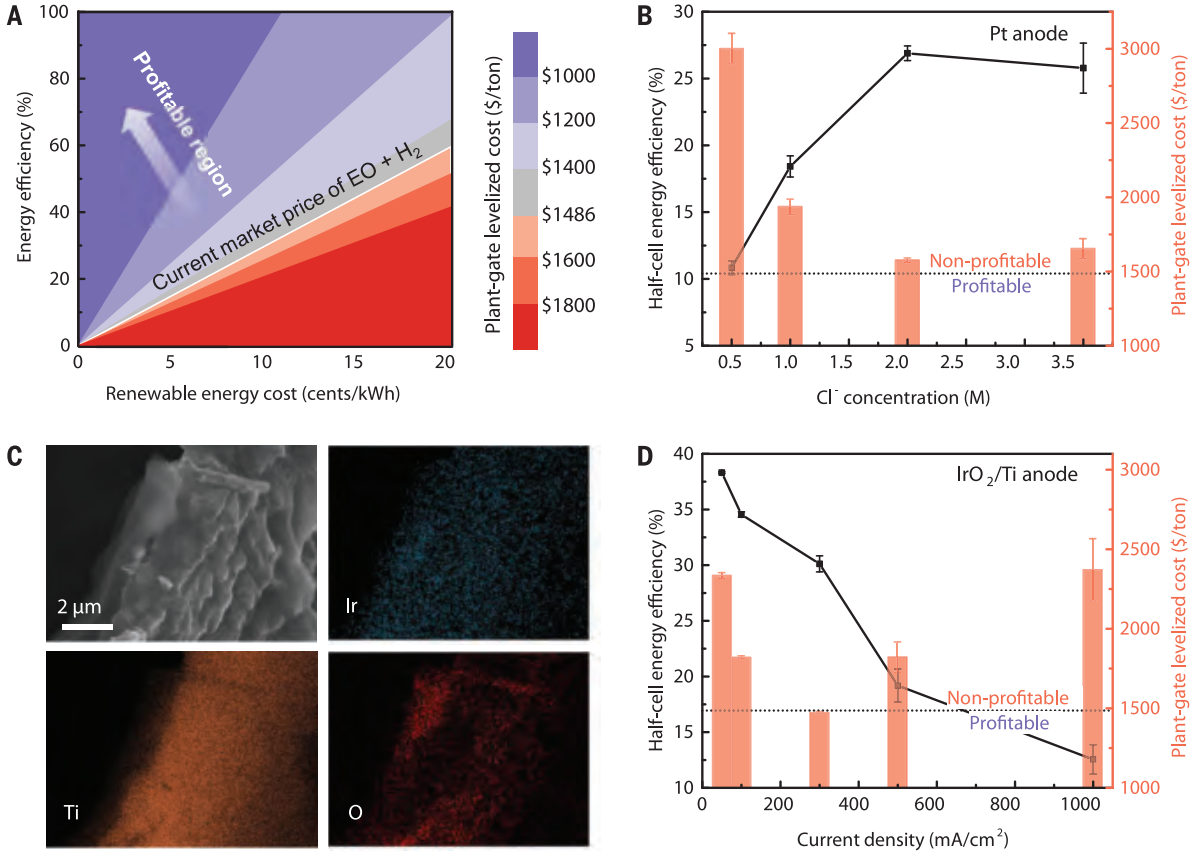


Fig. 3. Optimization of energy efficiency to reduce energy cost and maximize technoeconomic benefit. (**A**) TEA showing plant-gate levelized cost as a function of energy efficiency and renewable energy cost. (**B**) Half-cell energy efficiency and the corresponding plant-gate levelized cost as a function of Cl^- concentration. (**C**) EDX images showing the distribution of Ir, Ti, and O on the IrO_2/Ti mesh. (**D**) Half-cell energy efficiency and the corresponding plant-gate levelized cost as a function of current density. Our half-cell energy efficiencies are based on our reported potentials versus Ag/AgCl, which are not IR corrected. Additionally, we assume no losses at the cathode side, where hydrogen evolution occurs. The error bars correspond to the standard deviation of three independent measurements.

reaction (fig. S4) at the cathode during electrolysis generates the necessary OH^- (Eq. 4), whereas the AEM prevents complete mixing of the catholyte and the anolyte. Consequently, at the end of electrolysis, the pH of the catholyte becomes alkaline, with a pH value of 13.8. This means that by mixing the catholyte and anolyte output streams (performed after electrolysis), ethylene oxide can be generated from the reaction between ethylene chlorohydrin and OH^- (Eq. 5 and Fig. 2B). At the same time, the HCl generated (Eq. 2) is also neutralized by OH^- (Eq. 6).

The formation of ethylene chlorohydrin in the anolyte and subsequent generation of ethylene oxide in the mixing step were confirmed through ^1H nuclear magnetic resonance (NMR) (fig. S5). We performed the same experiments but using carbon-13–labeled ethylene ($^{13}C_2H_4$): ^{13}C NMR and ^1H NMR results confirm that the products we observed were indeed due to the partial oxidation of ethylene (Fig. 2C and fig. S5). In principle, a cation exchange membrane would be better in preventing crossover of OH^- (fig. S6A). However, this would lead to a continuous decrease in electrolyte (anolyte) conductivity during operation, resulting in lowered performance (fig. S6B). This system enables the generation of ethylene oxide in a single electrolyzer under ambient temperatures and pressures: Ethylene, water, and electricity are the consumables (Eq. 7).

Using this method, we achieved a Faradaic efficiency of 70(±1)% toward ethylene oxide (Fig. 2D) with 1.0 M KCl at 300 mA/cm^2. This corresponds to 3.9 mmol of ethylene oxide produced after 1 hour of electrolysis. Similar Faradaic efficiencies of 71(±1)% and 70(±1)% are maintained even at current densities of 500 and 800 mA/cm^2, respectively (Fig. 2D). A possible explanation for the missing charge could be O_2 evolution or complete oxidation of ethylene to form CO_2; however, when we performed gas chromatography on the output gas stream, we did not detect O_2 nor CO_2. The product specificity has a value of 97%; ethylene conversion to other products (such as CO_2) was not observed, and the remaining specificity was due to incomplete conversion of ethylene chlorohydrin to ethylene oxide (Eq. 5).

We instead hypothesized that the missing charge could be due to unreacted chlorine or hypochlorite species in the electrolyte; this was confirmed by using iodometric titration (fig. S7 and table S1). Pt and iridium oxide/titantium (IrO_2/Ti) anodes resist corrosion even at high chlorine concentrations during the chlor-alkali process (*31*). A continuous flow reactor process could be adopted (as opposed to the current batch reactor mode), in which portions of the electrolyte are periodically siphoned off for ethylene oxide extraction. In this case, the electrolyte would be continuously replenished by fresh or regenerated solution, preventing excessive buildup of these corrosive species.

The method could also be used for the epoxidations of other olefins; for example, when we replaced ethylene with propylene, Faradaic efficiencies were 69 to 71% toward propylene oxide—a commodity chemical with a market of 10 million tons per year in the plastics industry

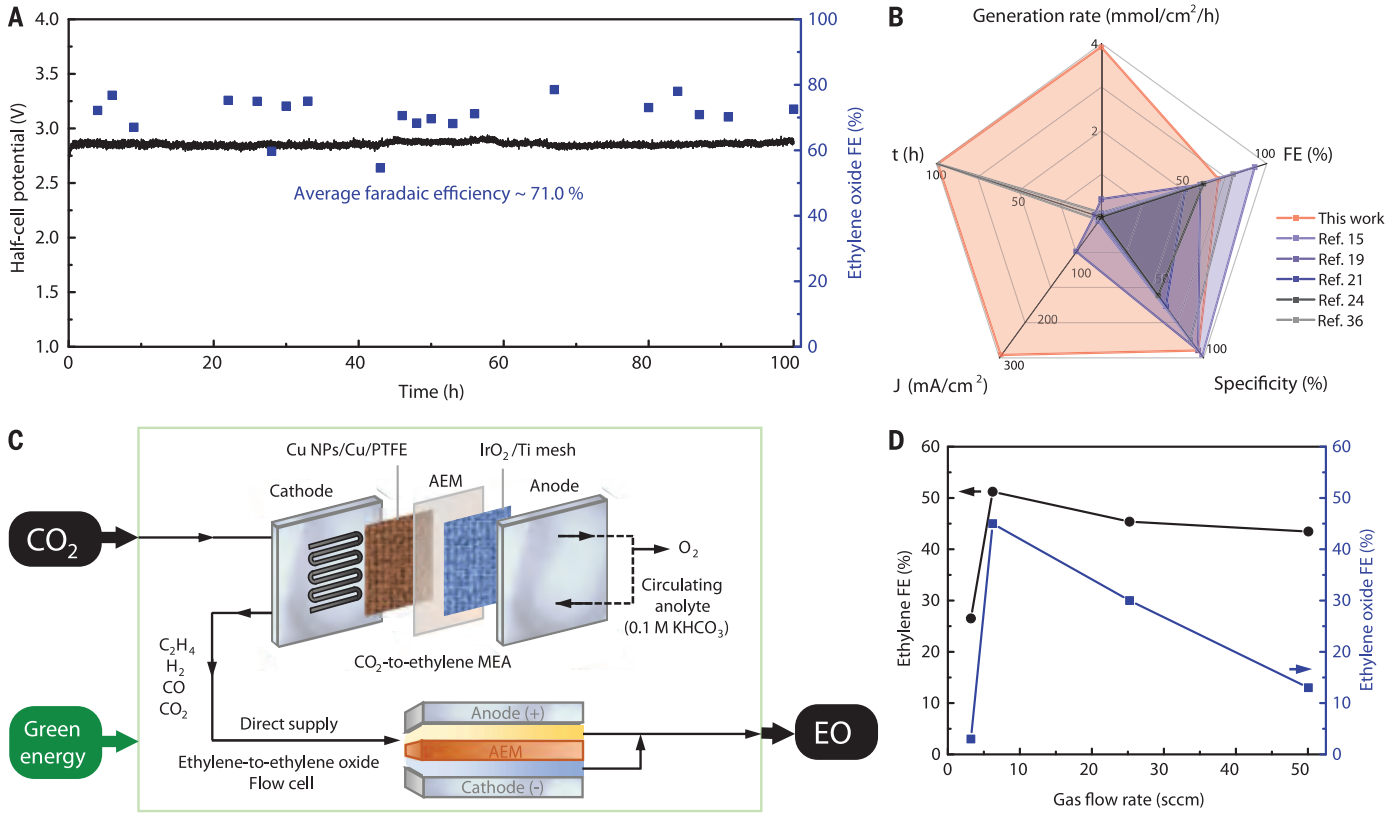


Fig. 4. Evaluation of ethylene-to–ethylene oxide performance. (**A**) Half-cell potential and Faradaic efficiency of ethylene oxide over 100 hours at 300 mA/cm^2. (**B**) Comparison of current density, product generation rate, reported operation time, Faradaic efficiency, and product specificity against state-of-the-art anodic upgrading reactions. Specificity refers to the percentage of all reacted substrate going toward the desired product. (**C**) Schematic of the CO_2-to–ethylene oxide (EO) process in which the ethylene-to-EO cell was directly supplied with the gas output from a CO_2-to-ethylene MEA. (**D**) Faradaic efficiencies of ethylene (in MEA) and ethylene oxide (in flow cell) as a function of the gas flow rate. For all cases, the MEA was run at 240 mA/cm^2, and the ethylene oxidation flow cell was operated at 300 mA/cm^2, for a duration of 1 hour.

(*32*)—at current densities of 300 to 800 mA/cm^2 (Fig. 2E).

We performed a technoeconomic analysis (TEA) to identify conditions that could enable the profitable synthesis of a renewable energy–powered anodic partial oxidation of ethylene to ethylene oxide (fig. S1) (full details of the TEA are available in the supplementary materials). For the TEA, we set a base electricity cost of 10 ¢/kWh, which is at least twice the average present-day industrial electricity cost (fig. S8) (*6*). Sensitivity analysis reveals that the greatest dependency of the plant-gate levelized cost is on electrochemical parameters such as current density and Faradaic efficiency (fig. S8; the range of values considered for each parameteris are provided in table S2). On the basis of the current market price per ton of ethylene oxide and the corresponding quantity of hydrogen produced at the cathode, we determined that for a current density of 300 mA/cm^2, the minimum energy efficiency required for the renewable energy–powered process to be profitable is ~30%. We also calculated the minimum energy efficiencies required to be profitable for different electricity costs up to 20 ¢/kWh, showing profitable regions as a function of energy efficiency and electricity cost (Fig. 3A). Similarly, this was also calculated for the electrosynthesis of propylene oxide from propylene (fig. S9).

The sensitivity analysis in fig. S8 revealed that the plant-gate levelized cost is sensitive to electrochemical parameters such as Faradaic efficiency and cell potential. To reduce energy cost, we sought to increase the energy efficiency of the reaction by varying the electrolyte concentration while operating at 300 mA/cm^2. We began at a lower Cl^- concentration (0.5 M); however, O_2 evolution from water then dominates the anodic reaction, resulting in a low Faradaic efficiency of 30(±1)% and energy efficiency of 11(±1)% (Fig. 3B). As the Cl^- concentration increases (1.0 and 2.0 M), the potential decreases [5.8(±0.2) V and 4.0(±0.1) V] because of improved Cl^- oxidation kinetics and increased electrolyte conductivity, leading to increased Faradaic efficiencies [70(±1)% and 67(±1)%] and half-cell energy efficiencies [18(±1)% and 27(±1)%]. At 3.5 M, however, the energy efficiency was unimproved at 26(±1.9)% because the reduced potential [3.6(±0.2) V] was negated by a slight decrease in Faradaic efficiency to 55(±1)%, likely because the increased Cl^- concentration is unfavorable for the disproportionation of Cl_2 into HOCl and HCl (Eq. 2). Thus, on the basis of the corresponding plant-gate levelized costs, we determined the optimal Cl^- concentration to be 2.0 M. In this work, all potentials are reported versus Ag/AgCl and are not solution resistance corrected.

Even at the optimal Cl^- concentration, the renewable electricity–based plant-gate levelized cost remains higher than the current market price per ton of ethylene oxide and the corresponding quantity of hydrogen (Fig. 3B). We turned to the working electrode (catalyst) as another degree of freedom to decrease the potential. We prepared IrO_2 deposited on Ti mesh (Fig. 3C and fig. S10) using a dip coating and thermal decomposition procedure (*33*). X-ray photoelectron spectroscopy (XPS) results confirmed the presence of Ir in a +4 oxidation state (fig. S10, A to C). Scanning

electron microscopy (SEM) images show the microscale mesh structure of the IrO_2-coated Ti mesh (fig. S10D). Energy-dispersive x-ray spectroscopy (EDX) confirmed the presence of Ir and O on the Ti mesh, indicating the loading of IrO_2 on Ti (Fig. 3C). X-ray diffraction (XRD) was also performed on the IrO_2 coating as well as the bare Ti mesh (fig. S10E). Additionally, transmission electron microscopy (TEM) images of the IrO_2 were acquired (fig. S10, F and G). Using this catalyst, we lowered the required applied potential from 3.4(±0.1) V to 3.0(±0.1) V, thus further raising the half-cell energy efficiency to 30(±1)% at 300 mA/cm^2.

Having optimized the electrochemical system, we measured the energy efficiencies and plant-gate levelized costs under different current densities to determine the most economical conditions for industrial manufacturing (Fig. 3D). Faradaic efficiencies were maintained even at a current density of 1 A/cm^2 [60(±4)%]. However, a much higher potential of 6.5(±0.5) V was required to drive the larger current, leading to a low half-cell energy efficiency [12(±1)%]. On the other hand, the half-cell energy efficiency is high at 38.3(±0.1)% under 50 mA/cm^2; thus, the electricity cost per ton of ethylene oxide is at the lowest. However, the high capital cost associated with electrolyzer surface area resulted in an uneconomical plant-gate levelized cost. The plant-gate levelized cost is the lowest at 300 mA/cm^2, with good energy efficiency of 30(±1)% and acceptably low capital costs.

On the basis of this analysis, we investigated the stability of the catalyst system at the most profitable current density of 300 mA/cm^2, during which portions of the electrolyte were periodically removed for analysis and replaced with fresh electrolyte. The system maintained a stable applied potential of 2.86(±0.02) V and Faradaic efficiency averaging 71(±0.6)% for 100 hours continuously (Fig. 4A). Post-reaction analysis of the anode through SEM and EDX revealed no obvious structural changes of the Ti mesh surface nor loss of IrO_2 (fig. S11). The method substantially outperforms other reported anodic upgrading reactions in current density, product generation rate, and reported operation time while maintaining high Faradaic efficiency and ethylene oxide specificity (Fig. 4B). The specificity in this case is 95%; we did not observe the conversion of ethylene to other products (such as CO_2). This high specificity is important in an industrial process because the ethylene will likely be continuously recirculated to maximize usage.

Last, we sought to develop an integrated system to perform the electrosynthesis of ethylene oxide from CO_2 (rather than ethylene) as the starting feedstock. This provides a route to directly use renewable electricity for recycling CO_2 into a valuable commodity chemical. In this integrated system, CO_2 reduction to ethylene is first performed by using a membrane electrode assembly (MEA) in a gas diffusion configuration, with O_2 evolution from water as the corresponding anodic reaction (Fig. 4C). The MEA comprises a copper nanoparticle/copper/polytetrafluoroethylene (Cu NPs/Cu/PTFE) cathode and an IrO_2/Ti mesh anode separated by an AEM, through which 0.1 M potassium bicarbonate ($KHCO_3$) anolyte was continuously circulated. The operating current density was kept at 240 mA/cm^2, and the ethylene Faradic efficiency was generally maintained at 43 to 52% (Fig. 4D). We measured the flow rate of the output gas using a flow meter at the cathode gas outlet, and the output gas was directly sparged into the anolyte of the ethylene-to–ethylene oxide flow cell (operated at 300 mA/cm^2) without further purification.

With this method, we achieved a Faradaic efficiency of 45% toward ethylene oxide under a gas flow rate of 6 standard cubic centimeters per minute (sccm) (Fig. 4D), despite the presence of other easily oxidizable gases such as H_2 and CO relative to ethylene (23% H_2, 12% CO, and 12% ethylene) (fig. S12 and tables S3 and S4). Oxidation of these gases requires direct contact with the anode, whereas ethylene oxidation is mediated by the extended heterogeneous:homogeneous interface and thus occurs in the bulk electrolyte at a much higher rate. The Faradaic efficiency toward ethylene oxide was reduced at a higher gas flow rate owing to lowered ethylene concentration in the MEA output stream (fig. S12 and tables S3 and S4). However, decreasing the flow rate even further (3 sccm) resulted in a lowered Faradaic efficiency toward ethylene in the MEA. This reduces the ethylene supply available for conversion in the flow cell, resulting in a drop in the Faradaic efficiency toward ethylene oxide. Thus, both concentration and molar quantity of ethylene in the MEA output stream are important determinants for the Faradaic efficiency toward ethylene oxide in the flow cell.

This demonstration shows the viability of an integrated system for complete CO_2-to–ethylene oxide conversion. Further improvements are expected by optimizing the ethylene Faradaic efficiency and single-pass conversion in the MEA.

REFERENCES AND NOTES

1. U.S. Energy Information Administration (EIA), "Energy use in industry," in *Use of Energy Explained* (EIA, 2019).
2. S. Chu, Y. Cui, N. Liu, *Nat. Mater.* **16**, 16–22 (2016).
3. Z. W. Seh *et al*., *Science* **355**, eaad4998 (2017).
4. J. Zheng, S. Suh, *Nat. Clim. Chang.* **9**, 374–378 (2019).
5. P. De Luna *et al*., *Science* **364**, eaav3506 (2019).
6. M. Jouny, W. Luc, F. Jiao, *Ind. Eng. Chem. Res.* **57**, 2165–2177 (2018).
7. C. Xia, Y. Xia, P. Zhu, L. Fan, H. Wang, *Science* **366**, 226–231 (2019).
8. R. F. Service, *Science* **365**, 1236–1239 (2019).
9. S. Rebsdat, D. Mayer, in *Ullmann's Encyclopedia of Industrial Chemistry*. (Wiley, 2001).
10. World Petrochemicals Program (WP), "Ethylene," WP Report (SRI Consulting, 2009).
11. World Petrochemicals Program (WP), "Ethylene oxide," WP Report (SRI Consulting, 2009).
12. A. Boulamanti, J. A. Moya, *Energy Efficiency and GHG Emissions: Prospective Scenarios for the Chemical and Petrochemical Industry* (Publications Office of the European Union, 2017).
13. S. T. Wismann *et al*., *Science* **364**, 756–759 (2019).
14. K. M. Van Geem, V. V. Galvita, G. B. Marin, *Science* **364**, 734–735 (2019).
15. H. G. Cha, K.-S. Choi, *Nat. Chem.* **7**, 328–333 (2015).
16. N. Jiang, B. You, R. Boonstra, I. M. Terrero Rodriguez, Y. Sun, *ACS Energy Lett.* **1**, 386–390 (2016).
17. B. You, N. Jiang, X. Liu, Y. Sun, *Angew. Chem. Int. Ed.* **55**, 9913–9917 (2016).
18. T. Li, Y. Cao, J. He, C. P. Berlinguette, *ACS Cent. Sci.* **3**, 778–783 (2017).
19. R. S. Sherbo, R. S. Delima, V. A. Chiykowski, B. P. MacLeod, C. P. Berlinguette, *Nat. Catal.* **1**, 501–507 (2018).
20. J. Zheng *et al*., *Adv. Funct. Mater.* **27**, 1704169 (2017).
21. D. Liu *et al*., *Nat. Commun.* **10**, 1779 (2019).
22. Y. Kwon, Y. Birdja, I. Spanos, P. Rodriguez, M. T. M. Koper, *ACS Catal.* **2**, 759–764 (2012).
23. C. Dai *et al*., *J. Catal.* **356**, 14–21 (2017).
24. A. Winiwarter *et al*., *Energy Environ. Sci.* **12**, 1055–1067 (2019).
25. M. Rafiee, K. C. Miles, S. S. Stahl, *J. Am. Chem. Soc.* **137**, 14751–14757 (2015).
26. E. J. Horn *et al*., *Nature* **533**, 77–81 (2016).
27. M. Rafiee, F. Wang, D. P. Hruszkewycz, S. S. Stahl, *J. Am. Chem. Soc.* **140**, 22–25 (2018).
28. C. L. McCABE, J. C. Warner, *J. Am. Chem. Soc.* **70**, 4031–4034 (1948).
29. M. Eigen, K. Kustin, *J. Am. Chem. Soc.* **84**, 1355–1361 (1962).
30. W. Tong *et al*., *Nat. Energy* (2020).
31. R. K. B. Karlsson, A. Cornell, *Chem. Rev.* **116**, 2982–3028 (2016).
32. "Market Analytics: Propylene Oxide - 2018," *Markets & Profitability* (Nexant, 2018).
33. W. Luc, J. Rosen, F. Jiao, *Catal. Today* **288**, 79–84 (2017).
34. Y. Huang, X. Chong, C. Liu, Y. Liang, B. Zhang, *Angew. Chem. Int. Ed.* **57**, 13163–13166 (2018).
35. C. Huang, Y. Huang, C. Liu, Y. Yu, B. Zhang, *Angew. Chem. Int. Ed.* **58**, 12014–12017 (2019).
36. Y. Lum *et al*., *Nat. Catal.* **3**, 14–22 (2020).

ACKNOWLEDGMENTS

We thank D. Kopilovic and R. Wolowiec for their kind technical assistance. **Funding:** This material is based on work supported by the Ontario Ministry of Colleges and Universities (grant ORF-RE08-034), Natural Sciences and Engineering Research Council (NSERC) of Canada (grant RGPIN-2017-06477), Canadian Institute for Advanced Research (CIFAR) (grant FS20-154 APPT.2378), and University of Toronto Connaught Fund (grant GC 2012-13). D.S. acknowledges the NSERC E. W. R. Steacie Memorial Fellowship. **Author contributions:** E.H.S. supervised the project. W.R.L, Y.L., and E.H.S. conceived the idea and designed the experiments. W.R.L and Y.L. carried out all the experimental work. A.O. fabricated the IrO_2-coated Ti mesh electrodes and assisted in the CO_2-to–ethylene oxide conversion experiments. Y.W. performed SEM measurements. D.-H.N. performed the XRD measurements. B.C. carried out the TEM measurements. J.W. carried out the XPS measurements. T.-T.Z., F.L., and D.S. contributed to data analysis and manuscript editing. W.R.L, Y.L., and E.H.S. cowrote the manuscript. All authors discussed the results and assisted during the manuscript preparation. **Competing interests:** W.R.L, Y.L., and E.H.S. have filed provisional patent application no. 63/002,653 regarding the electrosynthesis of oxiranes. **Data and materials availability:** All experimental data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1228/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S12
Tables S1 to S4
References (*37*–*39*)

14 October 2019; resubmitted 26 January 2020
Accepted 1 April 2020
10.1126/science.aaz8459

MESOSCOPIC PHYSICS

h/e oscillations in interlayer transport of delafossites

Carsten Putzke[1]*, Maja D. Bachmann[2,3], Philippa McGuinness[2,3], Elina Zhakina[2], Veronika Sunko[2,3], Marcin Konczykowski[4], Takashi Oka[2,5], Roderich Moessner[5], Ady Stern[6], Markus König[2], Seunghyun Khim[2], Andrew P. Mackenzie[2,3]*, Philip J.W. Moll[1]*

Microstructures can be carefully designed to reveal the quantum phase of the wave-like nature of electrons in a metal. Here, we report phase-coherent oscillations of out-of-plane magnetoresistance in the layered delafossites $PdCoO_2$ and $PtCoO_2$. The oscillation period is equivalent to that determined by the magnetic flux quantum, *h/e*, threading an area defined by the atomic interlayer separation and the sample width, where *h* is Planck's constant and *e* is the charge of an electron. The phase of the electron wave function appears robust over length scales exceeding 10 micrometers and persisting up to temperatures of T > 50 kelvin. We show that the experimental signal stems from a periodic field modulation of the out-of-plane hopping. These results demonstrate extraordinary single-particle quantum coherence lengths in delafossites.

Electrons in vacuum carry the characteristics of both particles and waves, which has been demonstrated in interference experiments directly probing the phase information (*1*). In metals, the transport properties are usually well captured by the particle nature of the electron only, as described by the semiclassical Boltzmann equation. The wave-like character is masked by the high density of electrons and their interaction with the ionic lattice, which leads to a loss of the phase information in bulk phenomena. With experimental effort, samples can be fabricated on the mesoscopic length scale over which the phase of the electron is preserved, thus becoming observable in electronic transport. A well-known example is the Aharonov-Bohm effect (ABE) in nanoscopic rings of gold (*2*, *3*), which presents a solid-state analog of the interference experiment by Davisson and Germer (*1*). Common to these experiments is the creation of an artificial loop enclosing magnetic flux, which acts as a beam splitter.

Here, we report an unexpected robust manifestation of phase coherence intrinsic to the out-of-plane transport in single bars of the ultrapure delafossites $PdCoO_2$ and $PtCoO_2$. These materials are composed of highly conducting Pd/Pt layers separated by CoO_2 layers, resulting in a large transport anisotropy ρ_c/ρ_a exceeding 1000. The layered triangular crystal lattice leads to an almost hexagonal Fermi surface (FS) (*4*) with little warping, which has been well characterized by de Haas–van Alphen oscillations (*5*) and angle-dependent magnetoresistance oscillations (*6*, *7*). These materials are the most conductive oxides known, with an in-plane transport mean free path (mfp) of >20 μm at low temperatures (*5*, *8*, *9*).

The strong anisotropy is also reflected in the growth of thin, plate-like crystals, a common property of layered materials. Although mesoscopic quantum phenomena are successfully probed in the plane of quasi–two-dimensional (2D) ultrapure metals, achieving such electrical transport perpendicular to the layers is challenging. We have overcome this difficulty by using focused ion beam (FIB) microstructuring techniques (*10*). Starting from as-grown crystals, we milled pillars along the c-axis, thereby restricting the in-plane dimensions to a few micrometers. A typical structure designed for four-point resistivity measurements is shown in Fig. 1A. Because the depth d and width w of the pillar are both well below the mfp, the system enters the ballistic transport regime in the plane.

We found an oscillatory magnetoresistance by applying an in-plane magnetic field at low temperatures (Fig. 1B). These oscillations are clearly visible in the raw data ($\Delta\rho_{osc}/\rho \sim 5\%$; Fig. 1). To perform further analysis, we focused on the second derivative of the magnetoresistance (Fig. 1C). The oscillations are periodic in magnetic field and their periodicity scales inversely with the width of the pillars over an order of magnitude, from 1.2 to 12 μm (Fig. 1D).

The observed periodicity matches very well with that expected for a magnetic flux quantum $\Phi_0 = h/e$, with the Planck constant h and electron charge e, threading through an area S enclosed by two adjacent Pd/Pt layers and the sample side walls (dashed line in Fig. 1D). This gives an area $S = w^*c/3$, where c denotes the crystallographic unit cell lattice constant ($PdCoO_2$: 1.774 nm; $PtCoO_2$: 1.781 nm) (*11*). Owing to the ABC stacking, the unit cell contains three Pd/Pt layers so the relevant height is $c/3$. Such oscillations of the magnetoresistance, periodic in Φ_0, demonstrate quantum transport of coherent electron waves spanning the width of the entire sample. Oscillations are readily observed in samples as wide as 12 μm, indicating a macroscopic phase preservation in the metal. $PdCoO_2$ has been shown to have an extremely long ballistic mfp (*12*), but observation of phase coherence in a high-carrier-density metal over such a long distance is still unexpected. It is particularly noteworthy that no special care had to be taken to decouple the sample from the environment, such as ultralow temperatures or substrate decoupling [all samples are simply attached to a sapphire chip by epoxy glue; see the materials and methods (*13*)]. Nevertheless, all samples from different crystals showed a highly consistent picture of strong, long-ranged quantum coherence.

So far, we have only considered magnetic fields applied perpendicular to the sample surface. If indeed the oscillation frequency is set by the flux through the area S, then it would be natural to expect a sinusoidal dependence on the magnetic field angle when rotating within the Pd/Pt layer. The experimental frequency spectrum upon rotation is more complex, with multiple frequencies appearing (Fig. 2). A natural geometric interpretation of the angle dependence is found because of the FS topography of $PdCoO_2$. The almost perfect hexagonal FS, in contrast to a circular one, exhibits three preferential directions of electron motion perpendicular to the flat faces of the FS. In real space, this describes three interweaving subsystems of directional electron flow in the plane, each spanning its own area S_i (where i = 1...3; sketched in Fig. 2A). The flux enclosed in each subsystem contributes oscillations of frequency $|\vec{B} \cdot \vec{S}_i|/\Phi_0$ to the total conduction, leading to three branches in the frequency spectrum offset by 60° from each other. The difference in symmetry between the hexagonal FS and the rectangular sample shape divides the branches into two different types. The samples are cut such that one preferential direction of motion is aligned with a sample side wall. Therefore, one subsystem area is set by the full sample depth, whereas two symmetric branches are related geometrically to the sample width. The aspect ratio

[1]Laboratory of Quantum Materials (QMAT), Institute of Materials, École Polytechnique Fédéral de Lausanne (EPFL), 1015 Lausanne, Switzerland. [2]Max Planck Institute for Chemical Physics of Solids, 01187 Dresden, Germany. [3]School of Physics and Astronomy, University of St. Andrews, St. Andrews KY16 9SS, UK. [4]Laboratoire des Solides Irradiés, CEA/DRF/IRAMIS, École Polytechnique, CNRS, Institut Polytechnique de Paris, 91128 Palaiseau, France. [5]Max Planck Institute for the Physics of Complex Systems, 01187 Dresden, Germany. [6]Weizmann Institute of Science, Department of Condensed Matter Physics, Rehovot 76100, Israel.

*Corresponding author. Email: carsten.putzke@epfl.ch (C.P.); andy.mackenzie@cpfs.mpg.de (A.P.M.); philip.moll@epfl.ch (P.J.W.M.)

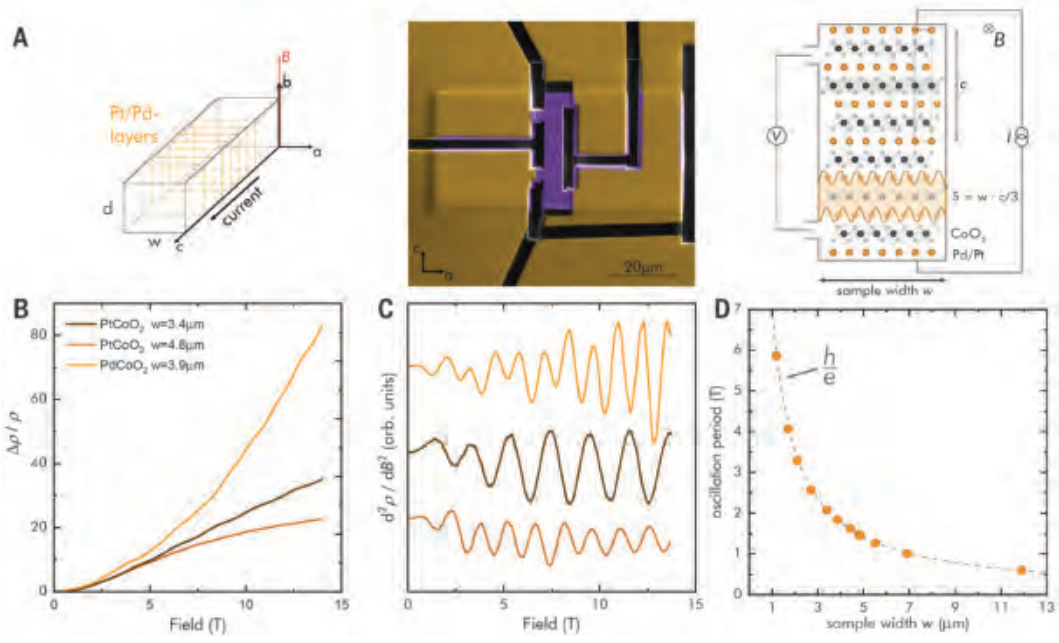


Fig. 1. Magnetoresistance oscillations periodic in magnetic field. (**A**) Experimental setup. (Left) Current is passed along a bar-shaped sample perpendicular to the layered structure. The bars have a width w and a thickness d. Magnetic field is applied and rotated within the Pd/Pt layer. (Middle) Scanning electron microscope image of $PdCoO_2$ microstructure to measure c-axis resistivity. (Right) Crystal structure of $PdCoO_2$ and $PtCoO_2$. Alternating layers of Pd/Pt and CoO_2 lead to a high anisotropy of the resistivity. The area S relevant for the h/e oscillations is spanned by two adjacent Pt/Pd layers. (**B**) Magnetoresistance of $PtCoO_2$ and $PdCoO_2$ at T = 2 K of various sample widths for fields along the a axis. The apparent difference in the high-field background is caused by a sharp feature in the angle-dependent magnetoresistance when fields are close to parallel with the Pd/Pt layers (*7*). (**C**) The second derivative of the resistivity highlights the oscillatory part of the magnetoresistance in (B). Multifrequency components are well explained by small sample misalignment (see Fig. 2). (**D**) The oscillation period is shown for different sample widths. The sample width dependence shows extraordinary agreement with the oscillation period expected for a single-particle magnetic flux quantum, h/e, per area $S = w*c/3$ (as indicated in the panel above).

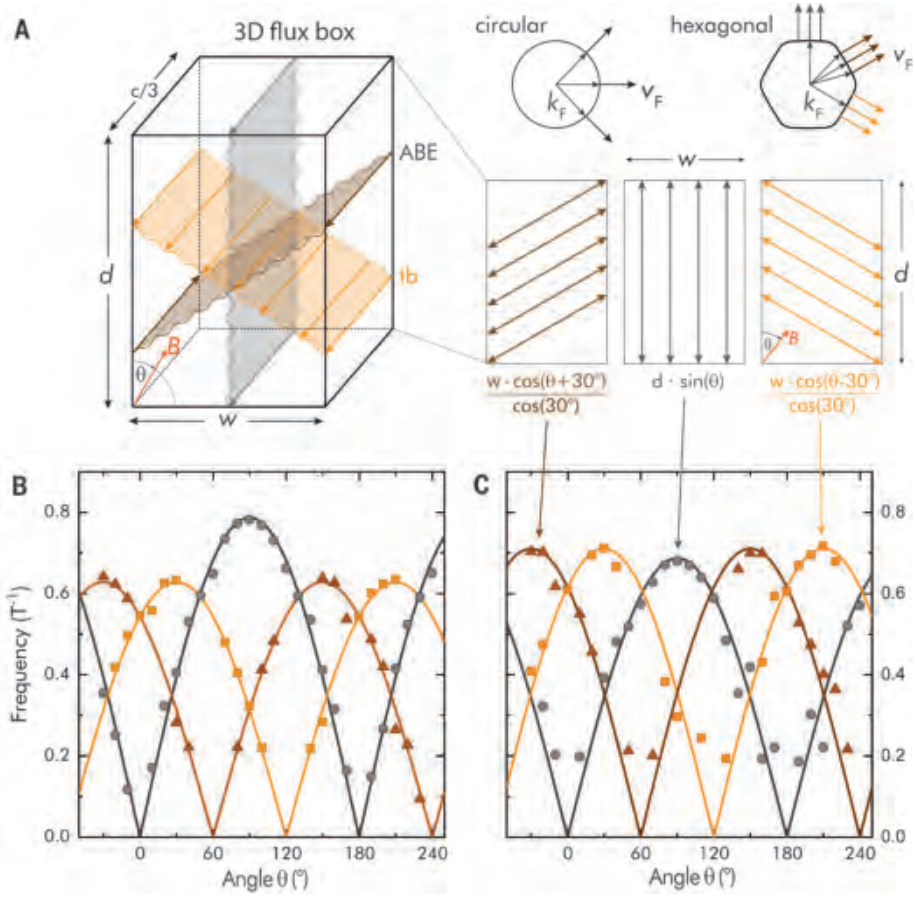


Fig. 2. Angle dependence of quantum-coherent oscillations. (**A**) $Pd/PtCoO_2$ has almost hexagonal FS's (*4*, *32*). This leads to three preferred directions of motion, unlike the case with a circular FS. The magnetic field is rotated in the planes. The three ballistic paths and their angle-dependent projections on the magnetic field are shown. In rectangular samples, two symmetric branches are set by the sample width and one by the sample thickness d (see discussion in main text). The left panel illustrates the relevant 3D flux box limited by the sample width, thickness, and two adjacent Pd/Pt layers. This box defines the flux surfaces. The oscillations are periodic in integer flux quanta threading through them. (**B** and **C**) Angle dependence of the quantum-coherent oscillations of $PdCoO_2$ (B) and $PtCoO_2$ (C). Solid symbols represent the measured data points, and solid lines represent the expectation from the model sketched in (A). (B) shows data from a sample with a d/w ratio of 1.4, whereas the aspect ratio for the sample in (C) is close to 0.9.

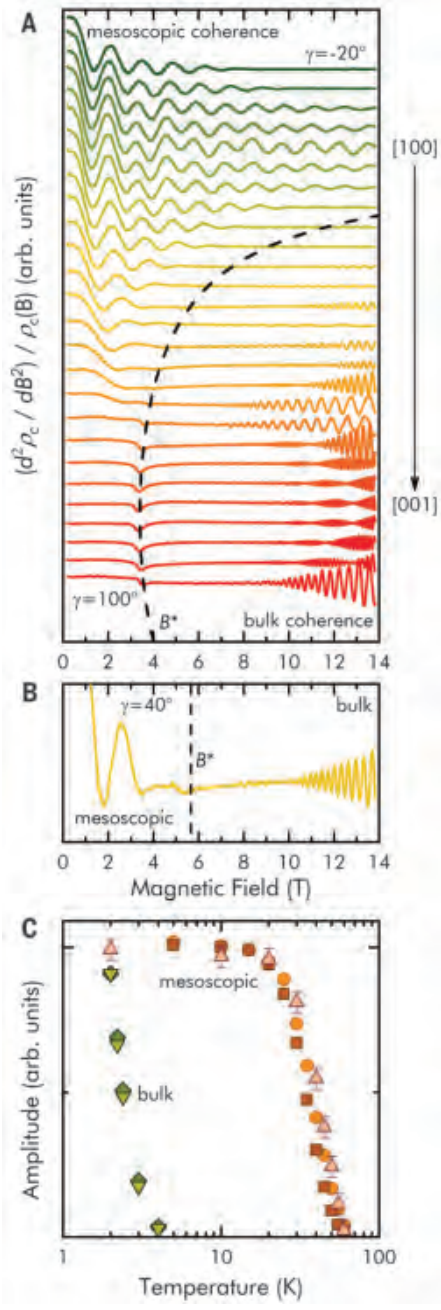


Fig. 3. Angle dependence tilting the field out of the plane. (**A**) Second derivative of the magnetoresistance with respect to the magnetic field. B-periodic oscillations in agreement with Fig. 1 for angle γ = 0° are shown. As the field is tilted out of the Pd layers in 5° steps, the oscillation period is modified as 1/cos γ. At higher tilt angles, the B-periodic oscillations vanish and SdH oscillations are observed (for a detailed analysis, see fig. S5). The dashed line represents the field B* at which the cyclotron diameter coincides with the sample width *w*. B-periodic oscillations are seen over a wide-angle range below B*, whereas SdH oscillations only appear above B*. The data have been offset proportional to the magnetic field angle. (**B**) Subset of the data in (A) at a magnetic field angle of γ = 40°. At low field, the B-periodic oscillations are seen, whereas at high field, 1/B-periodic oscillations are observed. (**C**) Mesoscopic oscillation amplitudes extracted from fast Fourier transform (FFT) analysis (see raw data in fig. S4) in the field range from 3 to 12 T and γ = 0° are shown for $PtCoO_2$ (*w* = 4.8 μm, squares; *w* = 2.0 μm, circles) and two samples of $PdCoO_2$ (#1: mfp = 20 μm, *w* = 3.9 μm, diamonds; #2: mfp = 1 μm, *w* = 1 μm, triangles). Despite the large changes of width and mfp, the temperature dependence of the signals from the different samples is very similar and the oscillations can be observed to temperatures of >50 K. By comparison, the temperature dependence of the conventional bulk SdH oscillations (γ = 90°) is shown for the two main orbits (see fig. S5).

of the cross section is reflected in the relative ratio of the maximum frequency values in the two types of branches. From the in-plane angle dependence, it follows that for the magnetic field aligned with both the sample sides (θ = 0° and 90°), an area S_i scales with *w* or *d*, respectively. In Fig. 1D, the data for both angle configurations were combined by denoting the dimension perpendicular to magnetic field as sample width *w*.

The period of all field-induced oscillations in quantum objects is expected to be the flux quantum threading through them, $B_n \cdot S = n\Phi_0$. Usually, the relevant length scale in metallic systems is set by the magnetic field itself, in the form of the cyclotron radius r_c, leading to oscillations periodic in $1/B$ ($B_n \cdot r_c^2 \propto n\Phi_0$). The best known of such $1/B$ periodic magnetoresistance oscillations are Shubnikov-de Haas (SdH) oscillations (*14*). Given the 2D nature of the FS in our system, SdH oscillations will not appear for in-plane fields because all orbits are open. However, as the magnetic field is rotated out of the Pd layers, the out-of-plane field induces orbital motion and usual SdH oscillations will appear. This leads to a particularly rich interplay between the different quantum transport regimes as a function of out-of-plane angle γ (Fig. 3). For an in-plane magnetic field (γ = 0°), *B*-periodic oscillations are observed as previously discussed. On tilting the field out of plane, those oscillations are limited to lower magnetic fields and vanish at a field scale *B** defined by the mesoscopic size of our samples (Fig. 3B). *B** corresponds to the angle-dependent field scale required to fit a bulk-like cyclotron radius into the pillar, given by the condition of $2r_c = w$. *B** appears as a clear anomaly in the magnetoresistance, delineating a strong negative magnetoresistance above *B** (fig. S3). Thereby, tilted magnetic fields induce a transition between mesoscopic quantum transport in the low-field regime and bulk-like transport described by Landau levels at sufficiently high fields. This picture is straightforward to understand: Once the in-plane Lorentz force is sufficient to bend a wave front back on itself, it will self-interfere, leading to Landau quantization. This detaches the wave function from the boundary, and thus the Landau levels are entirely bulk like and independent of sample width. This scenario is further supported by the negative magnetoresistance above *B** (fig. S3), which is caused by the suppression of the dominant boundary scattering in the clean devices owing to bulk-like Landau tubes forming in the core of the pillar. Indeed, above *B**, conventional SdH oscillations are observed in our samples, which coexist with *B*-periodic oscillations in the intermediate angle range (Fig. 3B). The SdH frequencies and effective masses of $m^* \approx 1.5\ m_e$ [$PdCoO_2$ (*5*)] and $1.2\ m_e$ [$PtCoO_2$ (*7*)] are consistent with work performed on macroscopic crystals, clearly excluding the possibility that the microfabrication has strongly altered the material. The large size of the hexagonal FS leads to high-frequency oscillations around F ~ 30 kT observed in the microstructure. However, resolving such high frequencies requires very slow field sweeps (<10 mT/min), which are impractical to perform over large field ranges and multiple angles. Therefore, only the slow difference frequency corresponding to the beating of neck and belly frequencies is apparent in Fig. 3; the main frequencies were always observed consistently when sweeping more slowly (fig. S5).

Further insights into the quantum transport arise from a comparison of the transport mfp and the quantum mfp. The quantum coherence length extracted from SdH oscillations is found to be only 400 nm (fig. S6), more than an order of magnitude smaller than that observed in the *B*-periodic oscillations. Furthermore, the quantum coherence length remains almost unchanged in irradiated samples, whereas the in-plane transport mfp is reduced by more than a factor of 10. It is important to consider that the quantum mfp obtained from a Dingle analysis

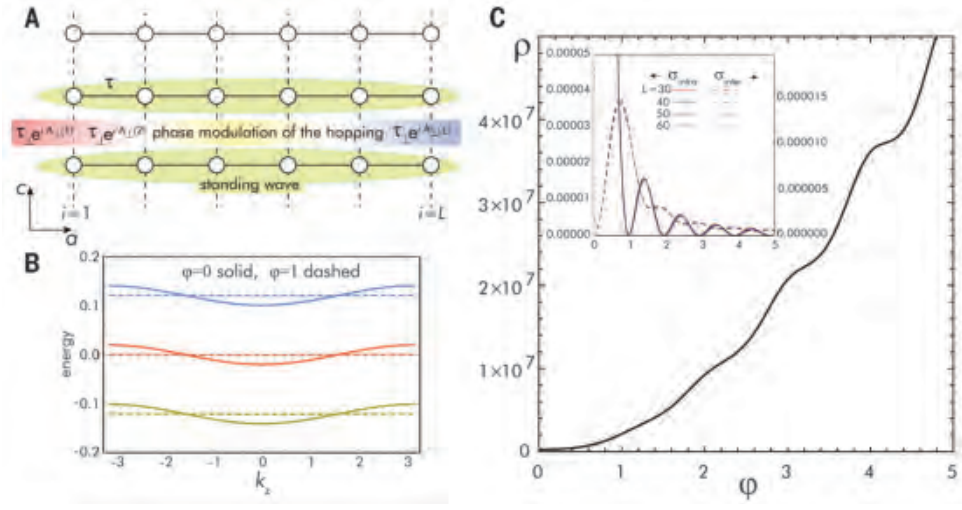


Fig. 4. Transport analysis for the anisotropic oHofstadter model. (**A**) Anisotropic Hofstadter model with tunneling parameters $\tau_1 = 100\tau_\perp = 1$ as a toy model to study the origin of the *h/e* magneto-oscillation shown in Fig. 1. Interlayer transport occurs as an electron in one layer, in an extended standing wave state, tunnels to an adjacent layer. In the presence of a magnetic field, the tunneling matrix element is modulated by a phase factor $e^{\imath 2\pi\frac{\varphi}{L}j}$ caused by the c-direction gauge field $A(j) = 2\pi\frac{\varphi}{L}j$, where φ is the flux per area between adjacent layers. This model naturally yields an effective interplane bandwidth given by $2\tau_\perp ReA(\varphi)$ in the large *L* and an anisotropic limit that vanishes for integer values of φ, where the factor $A(\varphi)$ resembling a diffraction grating effect is explained in the main text. (**B**) Energy spectrum of the anisotropic Hofstadter model around zero energy for $L = 51$ showing vanishing bandwidth at $\varphi = n$ (integer). (**C**) Resulting calculations of the interlayer resistivity. Full details of the calculation are presented in the materials and methods (*13*). In the absence of incoherent (e.g., phonon-assisted) interlayer processes, the resistivity would diverge for integer φ because only the intraband term σ_{intra} in the Kubo formula contributes (inset). However, for large *L*, the level separation shown in (B) becomes smaller than the level broadening from such incoherent processes, motivating the inclusion of interband contributions σ_{inter} in the calculation for the results shown in (C).

represents an average over the entire FS orbit, whereas the *B*-periodic oscillations stem exclusively from the flat sections of the hexagonal FS. A resolution to this conundrum would be a large quantum scattering rate at the corners of the hexagon; evidence of this has been reported previously in an analysis of the Hall effect (*15*).

B-periodic oscillations in a metal, such as those reported here, arise when a field-independent area *S* enters the quantization condition, such that $B_n \bullet S = n\Phi_0$. This most notably occurs in samples that are not simply connected geometrically. In these samples, e.g., Aharonov-Bohm rings or cylinders, a physical hole defines the relevant area. Such a physical area is naturally absent in our simply connected, bar-shaped samples. The experimental situation is also far from the so-called ultraquantum limit, at which the magnetic flux per crystalline unit cell is comparable to one flux quantum. This extreme limit requires fields on the order of $\Phi_0/\text{Å}^2 \sim 10^5$ T, at which *B*-linear oscillations from an atomic-scale flux box might be expected, yet such fields are inaccessible. The few known *B*-linear oscillatory phenomena in singly connected solids are semiclassical [Sondheimer resonances (*16*), Azbel-Kaner cyclotron motion (*17*), or geometric resonances in the presence of acoustic waves]; rely on a superconducting order parameter to establish macroscopic phase coherence [Fraunhoffer interference in Josephson junctions (*18*)] or exploit artificially introduced nanometric length scales or interference effects in tunneling between parallel quantum wires or wells (*19*). None of these can explain our data because the observed *h/e* periodicity clearly indicates long-range, single-particle-phase coherence as their origin.

Given the flux quantization condition that we have identified, it is at first sight appealing to invoke a scenario akin to the ABE to account for our data. In this picture, the quasiparticle would encircle the area *S* with the Pd/Pt layers resembling the arms of an interferometer. Because the nanoscopic dimensions of the interlayer distance would be combined with the macroscopic sample width, the necessary fields would be scaled to the range accessible in superconducting magnets. Although this would naturally lead to the observed periodicity, this scenario has severe shortcomings. First, the Pd layers are too strongly coupled through the sizable out-of-plane hopping element $\tau_\perp$ as determined by quantum oscillations ($\tau_1 = 1$ eV; $\tau_\perp = 10$ meV) (*8*), leading to an infinite number of paths involving a large number of layers. Furthermore, the ABE is most commonly accompanied by a related self-interference effect caused by weak localization called the Al'tshuler-Aronov-Spivak effect (*20*), which gives oscillations periodic in *h/2e* in metallic rings (*2*, *3*, *21*). This, or any other higher-order quantum process involving multiple layers in a stack (2*S*, 3*S*, …), would lead to higher harmonic content of the oscillations, which was not observed experimentally here within the noise level of 1% (fig. S2). A third key feature of our observation is its robustness to temperature (Fig. 3C). The observation of the ABE in metallic rings was limited to <4 K because of decoherence from various types of interaction with the environment, something that is also a feature of reported ABE experiments on graphene (*22*). This is in stark contrast to the $T > 50$ K temperature scale observed here in highly metallic $PdCoO_2$ and $PtCoO_2$. Similar high-temperature quantum coherence has been seen in bismuth nanowires (*23*), topological nanoribbons (*24*, *25*), quantum dots (*26*), carbon nanotubes (*27*), and 1/*B* superlattice oscillations in graphene (*28*), but the key to its observation was the small length scale involved in those nanoscale systems. This is not surprising in light of the calculation presented in the materials and methods (*13*), in which we show that even if we impose a far larger $\tau_\perp$ at the sample edge than in the bulk, an ABE-style calculation predicts an experimental signal with an amplitude dying out as 1/*w*. For our situation, this signal would be unresolvable.

Because the mfp decreases with increasing temperature, one would expect the *h/e* oscillations to persist up to higher temperatures for samples with smaller width *w*. By contrast, the onset temperature appears independent of sample width, which suggests that the upper temperature limit is not set by the mfp in the sample. To further probe this observation, we performed an additional set of measurements on a device in which the mfp had been reduced by a factor of 20 to 1 μm by the introduction of point defects created by 2.5 MeV of electron irradiation (*29*). Oscillations were unresolvable when the device width was 8 μm

and reappeared when it was narrowed to 1 μm (fig. S7), consistent with the expectation that the mfp must be on the order of the sample width or larger for the signal to be seen. Unexpectedly, the temperature dependence of the oscillations in the disordered device was the same as that in the cleaner devices within experimental error (Fig. 3C). This strongly suggests that if the zero temperature in-plane mfp exceeds the sample width, then the oscillatory signal persists until k_BT (where k_B is the Boltzmann constant) becomes a substantial fraction of $\tau_\perp$ despite the mfp becoming smaller than the device width at elevated temperatures.

This leads to a different scenario in which the purity of the experimental system plays an intriguing dual role. First, it provides a wave function with full phase coherence across the length of the *ab* plane. Second, the microscopically regular structure of the essentially perfect delafossite crystals gives rise to a periodic array of tunneling paths between the layers in the *c* direction.

In our picture, a phase coherent wave in the *ab* plane is transmitted to an adjacent plane with tunneling matrix elements $\tau_\perp << \tau_\parallel$ in each unit cell of the lattice. For a field applied along *b*, the phase of the tunneling matrix elements at site $j = 1...L$ is modulated by a factor $e^{\imath\, 2\,\pi \frac{w}{L} j}$. The field strength is written such that there is a total flux $\varphi = |\vec{B} \cdot \vec{S}_i|/\Phi_0$ per layer across the system of width L (= w/lattice constant) in the *a* direction. Summing the resulting series yields the following equation:

$$A(\varphi) = \sum_{j=1}^{L} e^{\imath\, 2\,\pi \frac{w}{L} j} = \frac{1 - e^{\imath\, 2\,\pi\varphi}}{e^{-\imath\, 2\,\pi \frac{w}{L}} - 1}$$

The reader will recognize this as equivalent to the far-field diffraction pattern from a diffraction grating of finite width L illuminated by a coherent light field. The finite width of such a grating is modeled by multiplying the transmission function of an infinite grating with a boxcar function (also known as a top hat function), rect(L). Invoking the convolution theorem of the Fourier transformation, its far-field diffraction pattern becomes convolved with the Fourier transform of rect(L), namely sinc(L). This is, we believe, the key physics behind our observation because this convolution gives rapid oscillations at a frequency that depends on L but with a relative amplitude from one oscillation to the next that is approximately L independent. In our experiment, the applied field is small, such that $\varphi/L << 1$, and we observed only the first few oscillations of the sinc function.

Reflecting this intuitive picture, we performed a transport calculation using the finite system Kubo formula (*30*–*32*) in the anisotropic Hofstadter model, which is summarized in Fig. 4 and presented in detail in the materials and methods. The calculation reproduces the qualitative experimental signal very well, with the expected sinc function clearly evident in Fig. 4C. It also provides a further key insight: The bandwidth for interplane transport (Fig. 4B) is seen to be modulated by the field, even vanishing for special fields at which φ is an integer, even though the bare-hopping $\tau_\perp$ remains unchanged.

The extremely anisotropic nature of the area pierced by the flux quantum—truly microscopic in one direction, almost macroscopic in the other—extends the experimentally accessible regime of the famous Hofstadter Hamiltonian to a region in which the hopping terms in the plaquette are highly anisotropic.

A natural question is why these oscillations, seen so clearly in our raw data, have not been observed before. The answer is that, in addition to the rarity of crystalline perfection at the level found in the delafossites, only recent technological advances have enabled experimental investigations of this regime (*10*). The technical key to this observation is to reduce w to a few micrometers, as well as to shape the sample cross section into an ideal rectangle by FIB. This acts as a magnifying glass that allows us to venture deeper into the mechanism behind coherent interlayer transport.

We believe that the observations and analysis reported here will stimulate further experimental, theoretical, and technological research. The framework that we have presented for understanding the data invites refinement, and a similar physical picture might be developed from slightly different starting viewpoints. It may also be possible to extend the experiments to other high-purity layered compounds such as the ruthenates. Furthermore, the evolution of the signal could be studied by fabricating bilayer and few-layer thin films. Such thin films may also be used to explore technological possibilities. As quantum coherence emerges as its own subject in technology, it will be interesting to explore whether applications can exploit the rare macroscopic, single-particle phase coherence in the delafossites.

REFERENCES AND NOTES

1. C. J. Davisson, L. H. Germer, *Proc. Natl. Acad. Sci. U.S.A.* **14**, 317–322 (1928).
2. V. Chandrasekhar, M. J. Rooks, S. Wind, D. E. Probe, *Phys. Rev. Lett.* **55**, 1610–1613 (1985).
3. R. A. Webb, S. Washburn, C. P. Umbach, R. B. Laibowitz, *Phys. Rev. Lett.* **54**, 2696–2699 (1985).
4. H. J. Noh *et al*., *Phys. Rev. Lett.* **102**, 256404 (2009).
5. C. W. Hicks *et al*., *Phys. Rev. Lett.* **109**, 116401 (2012).
6. J. C. A. Prentice, A. I. Coldea, *Phys. Rev. B* **93**, 245105 (2016).
7. N. Kikugawa *et al*., *Nat. Commun.* **7**, 10903 (2016).
8. H. Takatsu *et al*., *Phys. Rev. Lett.* **111**, 056601 (2013).
9. A. P. Mackenzie, *Rep. Prog. Phys.* **80**, 032501 (2017).
10. P. J. W. Moll, *Annu. Rev. Condens. Matter Phys.* **9**, 147–162 (2018).
11. M. Tanaka, M. Hasegawa, H. Takei, *J. Cryst. Growth* **173**, 440–445 (1997).
12. M. D. Bachmann *et al*., *Nat. Commun.* **10**, 5081 (2019).
13. See the supplementary materials.
14. D. Shoenberg, *Magnetic Oscillations in Metals* (Cambridge Univ. Press, 1984).
15. N. Nandi *et al*., *npj Quantum Mater* **3**, 66 (2018).
16. E. H. Sondheimer, *Phys. Rev.* **80**, 401–406 (1950).
17. M. I. Azbel, E. A. Kaner, *Sov. Phys. JETP* **30**, 811 (1956).
18. S. Ooi, T. Mochiku, K. Hirata, *Phys. Rev. Lett.* **89**, 247002 (2002).
19. S. Datta *et al*., *Phys. Rev. Lett.* **55**, 2344–2347 (1985).
20. B. L. Al'tshuler, A. G. Aronov, B. Z. Spivak, *P. Zh. Eksp. Teor. Fiz* **33**, 101 (1981).
21. C. P. Umbach, S. Washburn, R. B. Laibowitz, R. A. Webb, *Phys. Rev. B* **30**, 4048–4051 (1984).
22. S. Russo *et al*., *Phys. Rev. B Condens. Matter Mater. Phys.* **77**, 085413 (2008).
23. A. Nikolaeva, D. Gitsu, L. Konopko, M. J. Graf, T. E. Huber, *Phys. Rev. B* **77**, 17–19 (2008).
24. H. Peng *et al*., *Nat. Mater.* **9**, 225–229 (2010).
25. L. X. Wang, C. Z. Li, D. P. Yu, Z. M. Liao, *Nat. Commun.* **7**, 10769 (2016).
26. I. R. Sellers, V. R. Whiteside, I. L. Kuskovsky, A. O. Govorov, B. D. McCombe, *Phys. Rev. Lett.* **100**, 136405 (2008).
27. A. Bachtold *et al*., *Nature* **397**, 673–675 (1999).
28. R. Krishna Kumar *et al*., *Science* **357**, 181–184 (2017).
29. V. Sunko *et al*., *Phys. Rev. X* **10**, 021018 (2020).
30. G. Czycholl, B. Kramer, *Solid State Commun.* **32**, 945–951 (1979).
31. Y. Imry, N. S. Shiren, *Phys. Rev. B* **33**, 7992–7997 (1986).
32. D. J. Thouless, S. Kirkpatrick, *J. Phys. C Solid State Phys.* **14**, 235–245 (1981).
33. C. Putzke *et al*., Data for: *h/e* oscillations in interlayer transport of delafossites, Zenodo (2020); https://doi.org/10.5281/zenodo.3725930.

ACKNOWLEDGMENTS

We thank B. Ramshaw, R. Ilan, M. Fischer, A. Morpurgo, J. Schmalian, and L. Forro for helpful discussions. **Funding:** This project was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant no. 715730, MiTopMat) and also was supported by the Max Planck Society. A.P.M. and R.M. acknowledge support from the Würzburg-Dresden Cluster of Excellence on Complexity and Topology in Quantum Matter (EXC 2147). M.D.B., P.M., and V.S. acknowledge studentship funding from the EPSRC under grant no. EP/L015110/1. A.S. was supported by the Israel Science Foundation, the European Research Council (Project LEGOTOP), and the DFG through project no. CRC-183. M.K. acknowledges support from the SIRIUS irradiation facility through project no. EMIR 2019 18-7099. **Author contributions:** C.P., M.D.B., P.M., E.Z., and M.K. fabricated the microstructures. V.S., E.Z., P.M., and M.K. performed the electron irradiation. S.K. grew the single crystals. C.P. and P.J.W.M. performed the transport measurements and data analysis. C.P., T.O., R.M., A.S., A.P.M, and P.J.W.M. interpreted the data. All authors were involved in designing the experiment and writing the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All experimental data in the main text and supplementary materials are available at Zenodo (*33*).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1234/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
Table S1
References (*34*–*37*)

24 July 2019; accepted 22 April 2020
10.1126/science.aay8413

ICE SHEETS

Pervasive ice sheet mass loss reflects competing ocean and atmosphere processes

Ben Smith[1]*, Helen A. Fricker[2], Alex S. Gardner[3], Brooke Medley[4], Johan Nilsson[3], Fernando S. Paolo[3], Nicholas Holschuh[5,6], Susheel Adusumilli[2], Kelly Brunt[7], Bea Csatho[8], Kaitlin Harbeck[9], Thorsten Markus[4], Thomas Neumann[4], Matthew R. Siegfried[10], H. Jay Zwally[4,7]

Quantifying changes in Earth's ice sheets and identifying the climate drivers are central to improving sea level projections. We provide unified estimates of grounded and floating ice mass change from 2003 to 2019 using NASA's Ice, Cloud and land Elevation Satellite (ICESat) and ICESat-2 satellite laser altimetry. Our data reveal patterns likely linked to competing climate processes: Ice loss from coastal Greenland (increased surface melt), Antarctic ice shelves (increased ocean melting), and Greenland and Antarctic outlet glaciers (dynamic response to ocean melting) was partially compensated by mass gains over ice sheet interiors (increased snow accumulation). Losses outpaced gains, with grounded-ice loss from Greenland (200 billion tonnes per year) and Antarctica (118 billion tonnes per year) contributing 14 millimeters to sea level. Mass lost from West Antarctica's ice shelves accounted for more than 30% of that region's total.

Observations of ice sheet mass change are essential to our understanding of present and future sea level change (*1*–*3*). Ice sheets gain mass through snow accumulation and lose it through three processes: surface melt runoff (Greenland, 50 to 65%), iceberg calving (Antarctica, ~50%, and Greenland, 15 to 25%), and basal melting of floating ice shelves (Antarctica, ~50%) and tidewater glaciers (Greenland, 15 to 25%) (*4*–*6*). The net balance between these competing processes largely dictates decadal to centennial ice sheet contributions to sea level and depends on interactions between ice, ocean, and atmosphere. Surface meltwater runoff, basal melting, and precipitation are all expected to increase in a warming climate, which has been observed for both ice sheets (*7*, *8*). In Greenland, with a sea level potential of ~7 m, enhanced surface melt has resulted in widespread thinning of the ablation zone (*9*), and thinning and retreat of tidewater glacier fronts have led to accelerated flow (*10*), increased discharge toward the ocean (*6*), and near-coastal thinning (*11*–*13*) owing to increased flux divergence. In Antarctica, with a sea level potential of ~58 m, changes in ocean heat content linked to changes in Southern Hemisphere atmospheric conditions (*14*) have enhanced basal melting of ice shelves, causing them to shrink (*15*, *16*), which has reduced their buttressing capability and has led to increased ice discharge into the ocean (*17*–*20*). Despite rapid advancement in our ability to observe the ice sheet response to climate change, observation of ice mass changes associated with atmospheric and ocean forcing of the ice sheets with a self-consistent data set has been challenging. We used measurements from NASA's Ice, Cloud and land Elevation Satellite (ICESat; 2003–2009) and ICESat-2 (2018–2019) missions to resolve precise patterns of ice sheet height change, which when combined with a new firn model provide a combined estimate of total grounded and floating mass change from both ice sheets. Although loss of floating ice makes no direct contribution to sea level, it directly affects the rate of ice flow into the ocean. Patterns of change in floating and grounded ice together reveal the spatial signatures of the atmospheric and ocean processes that lead to grounded ice loss.

Satellite radar and laser altimeters have collected nearly continuous measurements since the early 1990s, providing one of the longest records of ice sheet change and revealing broad patterns of mass change across both ice sheets (*21*, *22*). In Antarctica, grounded ice changes have been qualitatively linked to floating ice shelf changes, but altimeter studies have all considered grounded ice (*1*, *23*) and floating ice (*1*, *16*) separately. The resulting differences in instruments, methodologies, and study periods can obscure connections between processes in grounded and floating ice, so a unified estimate of, for example, the ratio between grounded and floating ice loss has not been straightforward. Compared with radar altimetry, laser altimetry has the advantage of definitive measurement of the ice sheet surface with minimal subsurface penetration and the capability for accurate measurements over the steeper sloping ice sheet margins. ICESat, Earth's first polar-orbiting satellite laser altimeter, sampled the surface with small (~60 m) footprints and fine sampling (172 m), but results from that mission alone span only the short (6 years) duration of the mission. Subtle

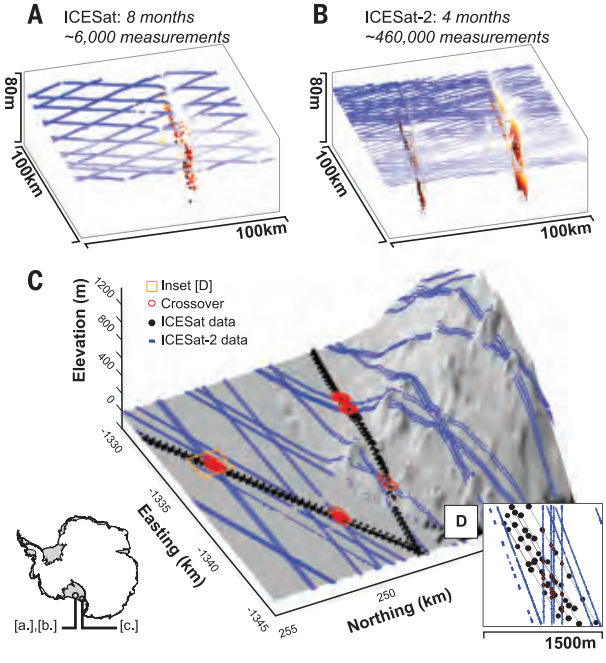


Fig. 1. Relative observation density of ICESat and ICESat-2 over the same target: Rifts of Ross Ice Shelf. (**A**) ICESat. (**B**) ICESat-2. Increased along-track resolution and cross-track observation density allow us to capture high-slope, small-scale features in unprecedented detail. (**C** and **D**) ICESat–ICESat-2 surface height comparison is done at the survey crossover points (red). ICESat-2's small footprint and dense along-track spacing [(D), to scale], combined with its repeat-track mission design, will result in the most precise measurements of height-change rates available to date.

[1]Polar Science Center, Applied Physics Laboratory, University of Washington, Seattle, WA, USA. [2]Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA. [3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. [4]Cryospheric Science Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA. [5]Department of Earth and Space Sciences, University of Washington, Seattle, WA, USA. [6]Department of Geology, Amherst College, Amherst, MA, USA. [7]Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA. [8]Department of Geological Sciences, University at Buffalo, Buffalo, NY, USA. [9]KBR, Greenbelt, MD, USA. [10]Department of Geophysics, Colorado School of Mines, Golden, CO, USA.
*Corresponding author. Email: besmith@uw.edu

changes in the ice sheet interiors have been difficult to capture because of measurement uncertainties [time-varying biases in radar altimetry (*2*) and ICESat (*24*)]. When integrated over the vast ice sheet area, these biases can overwhelm small, but important, changes in ice sheet mass balance from either trends in precipitation or long-term imbalance between ice flow, accumulation, and runoff (*1*).

NASA's next-generation laser altimeter ICESat-2 was designed to eliminate many of these problems. Launched 15 September 2018, ICESat-2's laser has a high sampling rate (0.7 m along-track), narrow footprint (~14.5 m), and near-global coverage (±88° latitude) repeating every 3 months, with a six-beam geometry that enables instantaneous cross-track slope determination (fig. S1). We compared ICESat-2 data (October 2018 to February 2019) with data from the ICESat mission (September 2003 to October 2008), which sampled a more coarsely spaced set of tracks to ±86° latitude (Fig. 1A) (*25*). Height-change estimates from these instruments cover all of Greenland and 95% of Antarctica. We removed the influence of local topography between missions by extracting height-difference measurements only at locations where the two sets of tracks cross (Fig. 1, C and D). Because both measurements come from laser altimeters, they are not strongly biased by subsurface scattering and retain their accuracy in sloping coastal regions. After aggregating the difference measurements into a regular grid, we estimated height-change rates and applied several corrections to obtain equivalent changes in mass, including a customized firn correction, state-of-the-art glacial isostatic adjustment, elastic compensation of Earth's surface, ocean tides, and inverse barometer effect (*25*). We restricted our ice shelf analysis to areas that were covered by ice shelves throughout both missions, so the ice shelf mass changes directly associated with changes in ice shelf extent are excluded from our estimates.

In Greenland, we found strong thinning that extends around the entire coastline (Fig. 2), which decreases inland, giving way to thickening at elevations between 2000 and 2500 m in western and southern Greenland and at elevations closer to 1500 m in the northeast. The largest thinning rates were between 4 and 6 m year^{-1} in Jakobshavn and Kangerlussuaq glaciers, whereas the largest inland thickening rates were less than 0.15 m year^{-1}. The total mass change rate for the ice sheet between 2003 and 2019 was −200 ± 12 Gt year^{-1}, with a basin-by-basin variation from −48 ± 4 Gt year^{-1} in the northwest to 2 ± 2 Gt year^{-1}

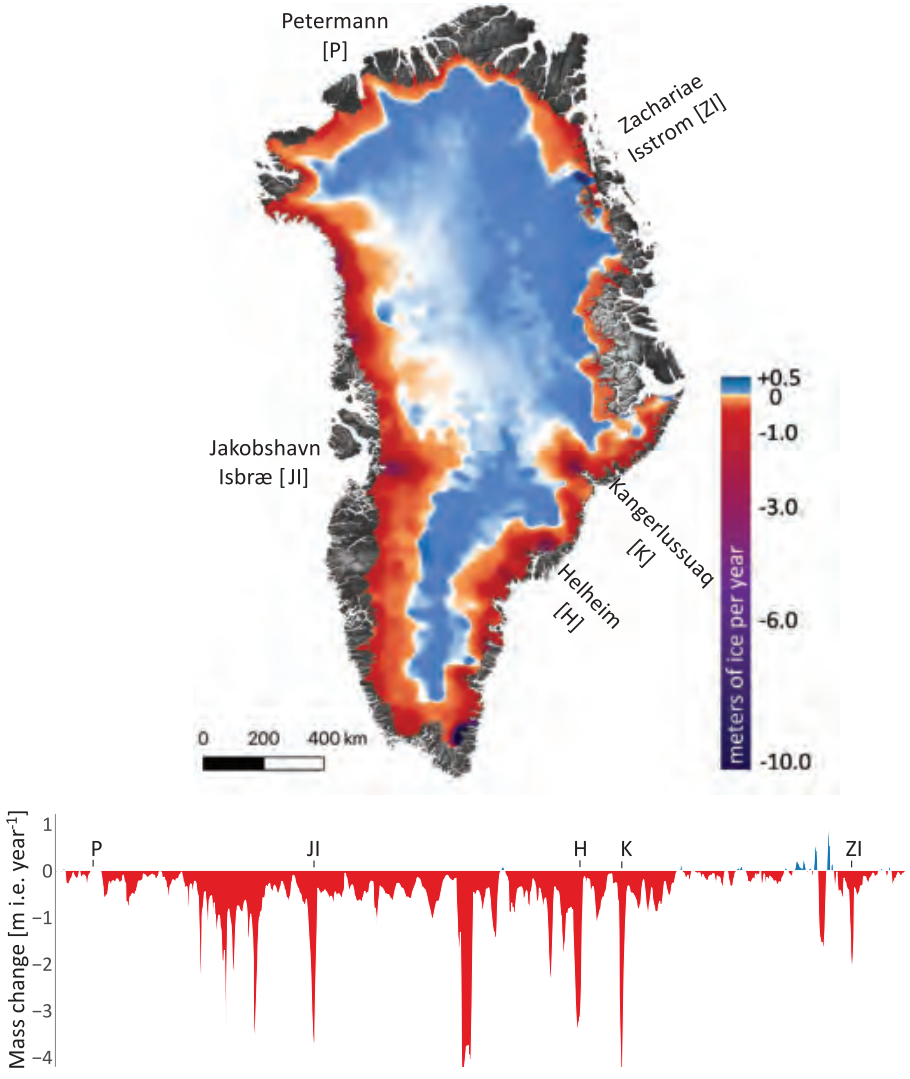


Fig. 2. Mass loss from Greenland Ice Sheet (2003 to 2019). (**Top**) Mass change for Greenland (meters of ice equivalent per year). (**Bottom**) Mass changes around the margin. Map and ice margin mass change have been smoothed with a 35-km median filter for improved visualization.

Table 1. Comparison of mass loss 2003 to 2019 for floating and grounded ice by region. SLE potential data are after (*44*): Greenland, East Antarctica (EAIS), West Antarctica (WAIS), and the Antarctic Peninsula (AP) (fig. S8). Cumulative ice loss and gain between 2003 and 2019 are provided in sea level equivalent (SLE) units.

	Change in mass over time (Gt year^{-1})		Sea level rise potential (m)	Total SLE 2003–2019 (mm)
	Floating ice	Grounded ice		
Greenland	N/A	−200 ± 12	7.4	8.9
EAIS	106 ± 29	90 ± 21	51.1	−4.0
WAIS	−76 ± 49	−169 ± 10	5.6	7.5
AP	−14 ± 28	−39 ± 5	0.5	1.7
Antarctica	15 ± 65	−118 ± 24	57.2	5.2

in the northeast (table S2). The low-elevation thinning is associated with both atmospheric and ocean processes: an increase in surface melt owing to a combination of increases in both air temperatures and exposure of bare ice during the summer (*26*–*28*). At the same time, the combination of increased surface melt and warmer ocean temperatures has led to enhanced submarine melting of submerged glacier termini (*29*, *30*) and has allowed more rapid calving by reducing the presence of rigid mélange in the fjords (*31*), each of which have increased glacier velocities and ice discharge into the ocean. With the exception of the northeast, every sector of the ice sheet lost substantial mass during our period of investigation.

Some of the highest Greenland ice mass losses are in the northwest and southeast sectors, where strong dynamic changes took place shortly after the start of the ICESat mission (*32*). The recent acceleration in ice loss from Northeast Greenland (*33*) appears more limited in extent and magnitude and has less impact on the total mass balance. Despite the record-setting discharge rates of Jakobshavn Isbrae (*34*), its contribution is only around 10% of the Greenland mass loss between 2003 and 2019, in part because the rapid mass loss due to its acceleration in the late 1990s (*35*) declined with the slowing and thickening of the lower part of the glacier between 2013 and 2018 (*36*). Overall, loss of solid ice around the margins outpaced lower rates of snow gain distributed across the interior.

In Antarctica, we see broad-scale patterns that are the fingerprints of two competing climate processes: snow accumulation and ocean melting. These processes occur on different spatial and temporal scales (Fig. 3) and exhibit strong connections between changes in grounded and floating ice in West Antarctica and the Antarctic Peninsula. The "background" pattern is one of subtle thickening along the steep slopes of the Antarctic Peninsula and around the coast to Queen Maud Land, East Antarctica, where gains decrease with distance from the ocean, which is indicative of snow accumulation in excess of that needed to balance flux divergence due to ice flow. This is likely due to enhanced moisture flux from marine air masses, but our measurements only provide an upper bound on the duration over which this may have occurred. Superimposed on this is a pattern of dramatic, ongoing mass loss around the margins, especially in the Amundsen and Bellingshausen regions of West Antarctica, which is likely in response to rapidly shrinking ice shelves. Ice shelf thinning in the Amundsen Sea has been attributed to an increase in atmospheric-driven incursions of modified Circumpolar Deep Water under the ice shelves, enhancing ocean-induced melting of marine-based basins (*14*, *16*). Similar patterns may be emerging for marine-based outlet glaciers of East Antarctica, such as at Denman Glacier (Fig. 3), where a deep subglacial canyon and a retrograde slope may drive unstable retreat (*37*). The three large cold-water ice shelves (Ross, Filchner-Ronne, and Amery) have smaller rates of height change, but there are striking internally driven changes where the stagnant Kamb Ice Stream (*38*) and slowing Whillans Ice Stream (*39*) starve downstream Ross Ice Shelf of mass input (locations are provided in fig. S8). In contrast to West Antarctic ice shelves, East Antarctic ice shelves gained 106 ± 29 Gt year^{-1} (Table 1).

The most substantial floating-ice losses occurred along the Amundsen-Bellingshausen region of West Antarctica and the Antarctic Peninsula. A basin-by-basin comparison between floating and grounded ice loss allows us to quantify the link between rapidly thinning ice shelves and grounded-ice loss in these regions; for example, 53% of mass loss from the Getz Ice Shelf basin (basin 20), 29% from

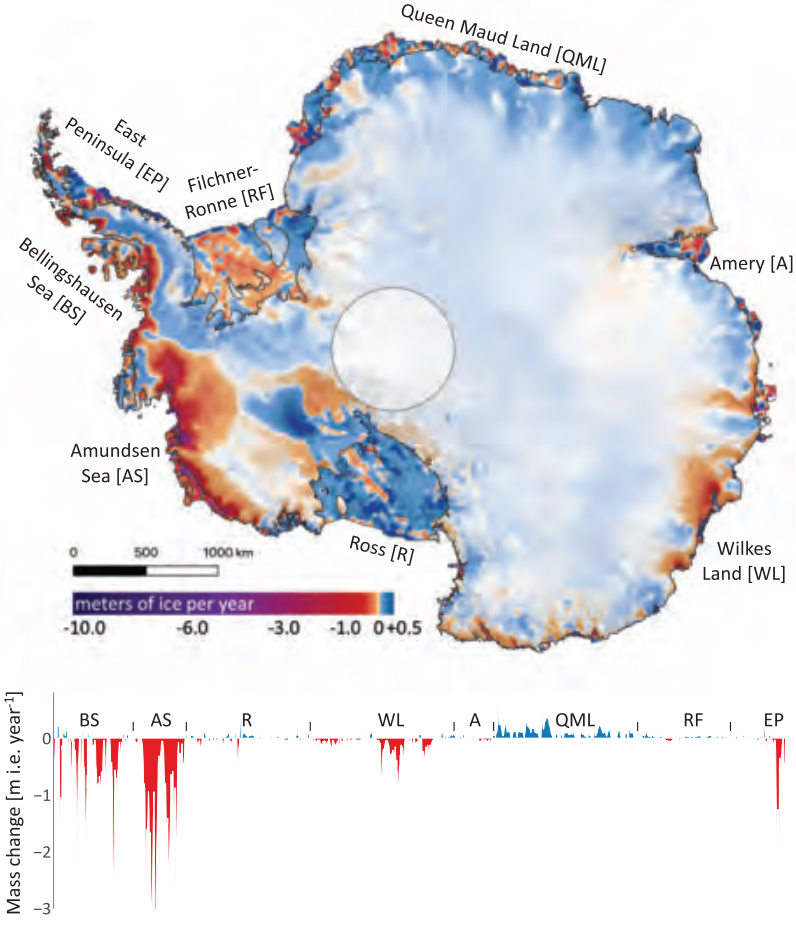


Fig. 3. Mass loss from Antarctica (2003 to 2019). (**Top**) Mass change for Antarctica. (**Bottom**) Mass changes at the grounding line. Highest mass loss rates are in West Antarctica and Wilkes Land, East Antarctica. Map and grounding line mass change have been smoothed with a 35-km median filter for improved visualization.

the Thwaites basin (basin 21), and 61% from the George VI Ice Shelf basin (basin 24) was from ice shelf mass loss. Overall, 31% of West Antarctic ice loss for 2003 to 2019 (76 ± 49 Gt $year^{-1}$) was from the ice shelves, whereas the remaining 69% (169 ± 10 Gt $year^{-1}$) was from the grounded ice feeding those ice shelves (Table 1 and table S1); for the Antarctic Peninsula, 27% of the mass loss was from floating ice.

Considering only grounded ice (for sea level contribution), our data show pervasive mass loss around the margins of Greenland, West Antarctica, and the Antarctic Peninsula, partially offset by mass gains in East Antarctica and central Greenland. Our mass loss rates of 118 ± 24 Gt $year^{-1}$ from Antarctica and 200 ± 12 Gt $year^{-1}$ from Greenland imply a total sea level contribution of 14 ± 1 mm over the 16-year period (Table 1). Compared with a compilation of mass-change estimates for a similar time span (2002 to 2017) (*2*), our Antarctic estimates are consistent (within reported errors) for the Antarctic Peninsula and for the whole ice sheet but significantly more positive for East Antarctica (90 ± 21 versus 2 ± 37 Gt $year^{-1}$) and significantly more negative for West Antarctica (–169 ± 10 versus –124 ± 27 Gt $year^{-1}$). For Greenland, our estimate is consistent with rates derived from a compilation of techniques that extends through 2018 (*40*) but is significantly more positive than some rates calculated from mass-flux techniques between 2003 and 2018 (–200 ± 12 versus –268 ± 14 Gt $year^{-1}$) (*41*). Another recent mass-flux–based estimate (*42*) gives modestly larger loss rates for Greenland (–233 + 12 Gt $year^{-1}$), with differences that arise in part because that estimate includes mass changes in peripheral ice and tundra that our study excludes. Because our estimates of height change have smaller instrumental biases than those of previous laser-altimetry estimates, our results suggest that earlier disagreement between some input-output and altimetry estimates (*1*, *8*) was at least partially due to negative biases in the input-output estimates. Despite this, our results show that the mass gains in East Antarctica are not sufficient to offset the rapid mass losses from West Antarctica and the Antarctic Peninsula, so that Antarctica's contribution to sea level change is unambiguously positive.

We estimated height changes for both ice sheets (grounded and floating ice) from NASA's ICESat and ICESat-2 laser altimetry missions (2003 to 2019). Applying new corrections to convert from height to mass change, we generated maps of mass changes for the ice sheets. Our maps highlight complex localized patterns of ocean-induced changes near the coast where inland ice is responding to increased frontal melt and ice shelf thinning by flowing faster, leading to increased flux divergence and surface lowering, with the strongest signals in the Amundsen and Bellingshausen coasts of Antarctica. Mixed signals, including the effects of surface melting and ocean-driven velocity changes, are apparent around the coasts of Greenland. We also see more subtle thickening across the vast interiors of the ice sheets, likely in response to increased snowfall, where the precision of measurements of height change by previous altimeters limited inferences of mass change. For both ice sheets, these patterns result from the interplay between the ocean and atmosphere; ultimately, high-elevation gains are greatly outmatched by low-elevation dynamic losses, combined with enhanced surface melting in Greenland.

Our unified estimates of grounded and floating ice sheet mass change show that overall, Greenland lost 200 ± 12 Gt $year^{-1}$, while Antarctica lost a total of 103 Gt $year^{-1}$, with 118 ± 24 Gt $year^{-1}$ from grounded ice and a small net gain of 15 ± 65 Gt $year^{-1}$ from ice shelves. Together, the ice sheets contributed ~14 mm sea level equivalent to the global oceans over that 16-year period (8.9 mm from Greenland and 5.2 mm from Antarctica). In West Antarctica, ice loss from ice shelves (which does not directly contribute to sea level change) accounted for 31% of the total mass loss, and all West Antarctic marine basins with ice grounded below sea level (which are sensitive to flow instabilities and whose losses directly contribute to sea level change) are out of balance. Given the susceptibility of ice shelves and floating glacier termini to changing atmospheric and oceanic conditions, and of grounded ice to shrinking ice shelves (*43*), we can expect increasing contribution from both Greenland and Antarctica to sea level rise on relatively short (decadal to centennial) time scales.

REFERENCES AND NOTES

1. A. Shepherd *et al*., *Science* **338**, 1183–1189 (2012).
2. IMBIE team, *Nature* **558**, 219–222 (2018).
3. J. L. Bamber, R. M. Westaway, B. Marzeion, B. Wouters, *Environ. Res. Lett.* **13**, 63008 (2018).
4. D. I. Benn, T. Cowton, J. Todd, A. Luckman, *Curr. Clim. Change Rep.* **3**, 282–290 (2017).
5. M. A. Depoorter *et al*., *Nature* **502**, 89–92 (2013).
6. E. Rignot, S. Jacobs, J. Mouginot, B. Scheuchl, *Science* **341**, 266–270 (2013).
7. J. T. M. Lenaerts, B. Medley, M. R. van den Broeke, B. Wouters, *Rev. Geophys.* **57**, 376–420 (2019).
8. E. Rignot *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **116**, 1095–1103 (2019).
9. M. van den Broeke *et al*., *Science* **326**, 984–986 (2009).
10. I. Joughin *et al*., *J. Geophys. Res. Earth Surf* **117**, F02030 (2012).
11. B. M. Csatho *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18478–18483 (2014).
12. I. M. Howat, B. E. Smith, I. Joughin, T. A. Scambos, *Geophys. Res. Lett.* **35**, L17505 (2008).
13. N. Wilson, F. Straneo, P. Heimbach, *Cryosphere* **11**, 2773–2782 (2017).
14. M. Thoma, A. Jenkins, D. Holland, S. Jacobs, *Geophys. Res. Lett.* **35**, L18602 (2008).
15. H. D. Pritchard *et al*., *Nature* **484**, 502–505 (2012).
16. F. S. Paolo, H. A. Fricker, L. Padman, *Science* **348**, 327–331 (2015).
17. J. J. Fürst *et al*., *Nat. Clim. Chang.* **6**, 479–482 (2016).
18. T. K. Dupont, R. B. B. Alley, *Geophys. Res. Lett.* **32**, 1–4 (2005).
19. T. A. Scambos, J. A. Bohlander, C. A. Shuman, P. Skvarca, *Geophys. Res. Lett.* **31**, L18402 (2004).
20. E. Rignot, J. Mouginot, M. Morlighem, H. Seroussi, B. Scheuchl, *Geophys. Res. Lett.* **41**, 3502–3509 (2014).
21. D. J. Wingham, A. Shepherd, A. Muir, G. J. Marshall, *Philos. Trans. A Math. Phys. Eng. Sci.* **364**, 1627–1635 (2006).
22. H. J. Zwally *et al*., *J. Glaciol.* **51**, 509–527 (2005).
23. H. D. Pritchard, R. J. Arthern, D. G. Vaughan, L. A. Edwards, *Nature* **461**, 971–975 (2009).
24. A. A. Borsa, H. A. Fricker, K. M. Brunt, *IEEE Trans. Geosci. Remote Sens.* **57**, 6946–6959 (2019).
25. Materials and methods are available as supplementary materials.
26. J. C. Ryan *et al*., *Nat. Commun.* **9**, 1065 (2018).
27. M. R. van den Broeke *et al*., *Cryosphere* **10**, 1933–1946 (2016).
28. X. Fettweis *et al*., *Cryosphere* **11**, 1015–1033 (2017).
29. M. Wood *et al*., *Geophys. Res. Lett.* **45**, 8334–8342 (2018).
30. D. Slater *et al*., *Cryosphere* **13**, 2489–2509 (2019).
31. T. Moon, I. Joughin, B. Smith, *J. Geophys. Res. Earth Surf.* **120**, 818–833 (2015).
32. E. Rignot, P. Kanagaratnam, *Science* **311**, 986–990 (2006).
33. S. A. Khan *et al*., *Nat. Clim. Chang.* **4**, 292–299 (2014).
34. I. Joughin, B. E. Smith, D. E. Shean, D. Floricioiu, *Cryosphere* **8**, 209–214 (2014).
35. I. Joughin, W. Abdalati, M. Fahnestock, *Nature* **432**, 608–610 (2004).
36. A. Khazendar *et al*., *Nat. Geosci.* **12**, 277–283 (2019).
37. M. Morlighem *et al*., *Nat. Geosci.* **13**, 132–137 (2020).
38. G. A. Catania, T. A. Scambos, H. Conway, C. F. Raymond, *Geophys. Res. Lett.* **33**, L14502 (2006).
39. L. H. Beem *et al*., *J. Geophys. Res. Earth Surf.* **119**, 212–224 (2014).
40. IMBIE Team, *Nature* **579**, 233–239 (2020).
41. J. Mouginot *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **116**, 9239–9244 (2019).
42. M. D. King *et al*., *Cryosphere* **12**, 3813–3825 (2018).
43. G. H. Gudmundsson, F. S. Paolo, S. Adusumilli, H. A. Fricker, *Geophys. Res. Lett.* **46**, 13903–13909 (2019).
44. K. J. Tinto *et al*., *Nat. Geosci.* **12**, 441–449 (2019).

ACKNOWLEDGMENTS

We thank the ICESat-2 project and the ICESat-2 science team for enormous contributions to this project. We are also grateful to two anonymous reviewers for their constructive and insightful reviews. **Funding:** This work was funded by the NASA Cryospheric Sciences Program in support of the ICESat-2 mission under the following awards: NNX15AE15G (to B.S.); NNX15AC80G (to H.A.F.); NNX16AM01G (to N.H.); NNX17AI03G (to H.A.F.). A.S.G., J.N., and F.S.P. were funded by the NASA Cryospheric Sciences and NASA MEaSUREs programs through direct award to A.S.G. **Author contributions:** H.A.F., B.S., and A.S.G. conceptualized the study and, with F.S.P., wrote the majority of the main text. B.M. conducted firn and climate modeling. F.S.P., S.A., and J.N. performed altimetry data processing. N.H., F.S.P., J.N., and A.S.G. developed figures. K.B. provided insights into error models and calibrations. B.C., M.R.S., and H.J.Z. provided insights into altimetry and data analysis. T.N., T.M., and K.H. were responsible for the success of the ICESat-2 mission. All authors contributed to the writing and editing of the manuscript. **Competing interests:** The authors report no competing interests. **Data and materials availability:** ICESat data are available at https://nsidc.org/data/gla12, and the ICESat-2 data are available at https://nsidc.org/data/atl06. Mass-change grids are available from https://digital.lib.washington.edu/researchworks/handle/1773/45388 (accession no. 45388).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1239/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S8
Tables S1 and S2
References (*45*–*100*)

20 September 2019; accepted 13 April 2020
Published online 30 April 2020
10.1126/science.aaz5845

ECOLOGICAL ECONOMICS

Cascading social-ecological costs and benefits triggered by a recovering keystone predator

Edward J. Gregr[1,2]*, Villy Christensen[3], Linda Nichol[4], Rebecca G. Martone[1,5], Russell W. Markel[1,5], Jane C. Watson[6], Christopher D. G. Harley[3,7,8], Evgeny A. Pakhomov[3,8,9], Jonathan B. Shurin[10], Kai M. A. Chan[1]

Predator recovery often leads to ecosystem change that can trigger conflicts with more recently established human activities. In the eastern North Pacific, recovering sea otters are transforming coastal systems by reducing populations of benthic invertebrates and releasing kelp forests from grazing pressure. These changes threaten established shellfish fisheries and modify a variety of other ecosystem services. The diverse social and economic consequences of this trophic cascade are unknown, particularly across large regions. We developed and applied a trophic model to predict these impacts on four ecosystem services. Results suggest that sea otter presence yields 37% more total ecosystem biomass annually, increasing the value of finfish [+9.4 million Canadian dollars (CA$)], carbon sequestration (+2.2 million CA$), and ecotourism (+42.0 million CA$). To the extent that these benefits are realized, they will exceed the annual loss to invertebrate fisheries (−$7.3 million CA$). Recovery of keystone predators thus not only restores ecosystems but can also affect a range of social, economic, and ecological benefits for associated communities.

As keystone species, top predators can exert strong effects over the function, structure, and diversity of ecosystems (*1*). When these species recover after extirpation, they often reestablish top-down control (*2*) and shift the ecosystem closer to an unexploited state (*3*). This can disrupt social-ecological systems established during the species' absence and lead to conflict between the recovering predator and established human resource users (*4*). Given the widespread defaunation of natural systems (*2*), the societal conflicts arising from such rewilding efforts need to be acknowledged and quantified. However, despite numerous examples of such conflicts (*5*–*7*), the associated social, economic, and ecological changes are rarely documented or evaluated, making it challenging to manage and equitably mitigate impacts.

We demonstrate such an evaluation here by examining the transformation under way in the eastern North Pacific, where sea otters (*Enhydra lutris*), a marine keystone species (*8*), are recovering after near extinction by the maritime fur trade of the 18th and 19th centuries. As predators of invertebrates, in particular kelp-grazing sea urchins, sea otters release kelp from grazing pressure and promote the growth of kelp forests. This increases primary production, fixes free CO_2, and provides vertical habitat for other coastal species, particularly fish (e.g., rockfish, greenlings, and salmon).

This well-studied trophic cascade (*8*, *9*) is broadly seen as a conservation success story and case study in marine ecosystem restoration. However, sea otter recovery is unpopular in many coastal communities where sea otters compete strongly with humans for valuable invertebrates like crabs, clams, and urchins. This has led to conflict with established commercial and subsistence invertebrate fisheries across much of the reoccupied sea otter range. The scope of human-induced sea otter mortality is largely unknown, but may be a factor in slowing range expansion (Fig. 1). Although this conflict was anticipated (*10*, *11*) and reduced invertebrate catches are regularly reported by fishers, the associated costs and potential benefits of sea otter recovery have not been quantitatively assessed (*12*).

Understanding the costs and benefits arising from different ecosystem states is central to effective and equitable resource management. Accordingly, assessments of ecosystem services trade-offs are increasingly common (*13*). However, modeling the complexities of social-ecological systems requires many simplifying assumptions (*14*), which foreclose on our ability to comprehensively assess the full range of values that matter to people (*15*). Different representations are thus necessary for different applications. For example, the literature has focused largely on economic valuation of ecosystem services (*16*, *17*), whereas ecosystem-based management of fisheries has focused on ecological interactions and indicators related to fisheries (*18*, *19*), ecosystem health (*20*), and biodiversity (*21*, *22*). Calibrating relevant indicators with empirical data (*23*) at a scale that accurately represents the system of interest (*17*), articulating them in a way that is informative to management (*24*), and effectively communicating uncertainty (*14*, *25*) remain considerable challenges. Here, to support adaptive resource management, we translate local studies into a regional assessment of four diverse ecosystem services and propose an intuitive and comprehensive method for representing uncertainty. We examine whether sea otter–induced changes in finfish catch, carbon sequestration, and tourism offset the associated acute and contentious economic losses to invertebrate fisheries. These services are all closely linked to the sea otter–induced trophic cascade.

Although our empirical results represent one region, they are representative of these effects across the sea otter range, with some variability (see supplementary materials). More broadly, our interdisciplinary approach of translating field studies into economic value using integrated models, with defensible and intuitive treatment of uncertainty, is broadly relevant across many social and ecological contexts.

We take advantage of a natural experiment under way in Pacific Canada, where sea otters have been reoccupying large parts of their historical range for several decades [(*9*); Fig. 1]. Using a trophic model calibrated with local data, we estimate—with uncertainty—the regional change in biomass resulting from the transformation of an ecosystem without sea otters to one with sea otters present. We then estimate the potential change in value of the four ecosystem services using data on fisheries catch and landed value, tourism choices, carbon pricing, and estimates of trophic transfer efficiency. We also consider how this transformation influences less quantifiable benefits to the broader coastal ecosystem and the nonmonetary services provided to coastal communities. We examine the parametric uncertainty in both the trophic model and the translation of system biomass into economic benefits. Predictions of biomass change are presented showing the range of values under different parameterizations. Uncertainties in the dollar value of the four ecosystem services are presented with credibility estimates intended to show the range of defensible values for each service (see supplementary materials for details).

Our model reproduces observed aspects of the trophic cascade, including the decline of valuable invertebrate species such as geoduck clam, Dungeness crab, and sea urchin and increases in kelp abundance, primary production, and the biomass of lower trophic levels (Fig. 2). The aggregate change in predicted ecosystem

[1]Institute for Resources Environment, and Sustainability, University of British Columbia, 2202 Main Mall, Vancouver, BC V6T 1Z4, Canada. [2]SciTech Environmental Consulting, 2136 Napier St., Vancouver, BC V5L 2N9, Canada. [3]Institute for the Oceans and Fisheries, University of British Columbia, 2202 Main Mall, Vancouver, BC V6T 1Z4, Canada. [4]Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Rd., Nanaimo, BC V9T 6N7, Canada. [5]Outer Shores Expeditions, P.O. Box 361, Cobble Hill, BC V0R 1L0, Canada. [6]Biology Department, Vancouver Island University, 900 5th St. Nanaimo, BC V9R 5S5, Canada. [7]Department of Zoology, University of British Columbia, 6270 University Blvd., Vancouver, BC V6T 1Z4, Canada. [8]Hakai Institute, P.O. Box 309, Heriot Bay, BC V0P 1H0, Canada. [9]Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, 2207 Main Mall, Vancouver, BC V6T 1Z4, Canada. [10]Section of Ecology, Behavior and Evolution, University of California, San Diego, 9500 Gilman Dr. #0116, La Jolla, CA 92093, USA.
*Corresponding author. Email: ed@scitechconsulting.com

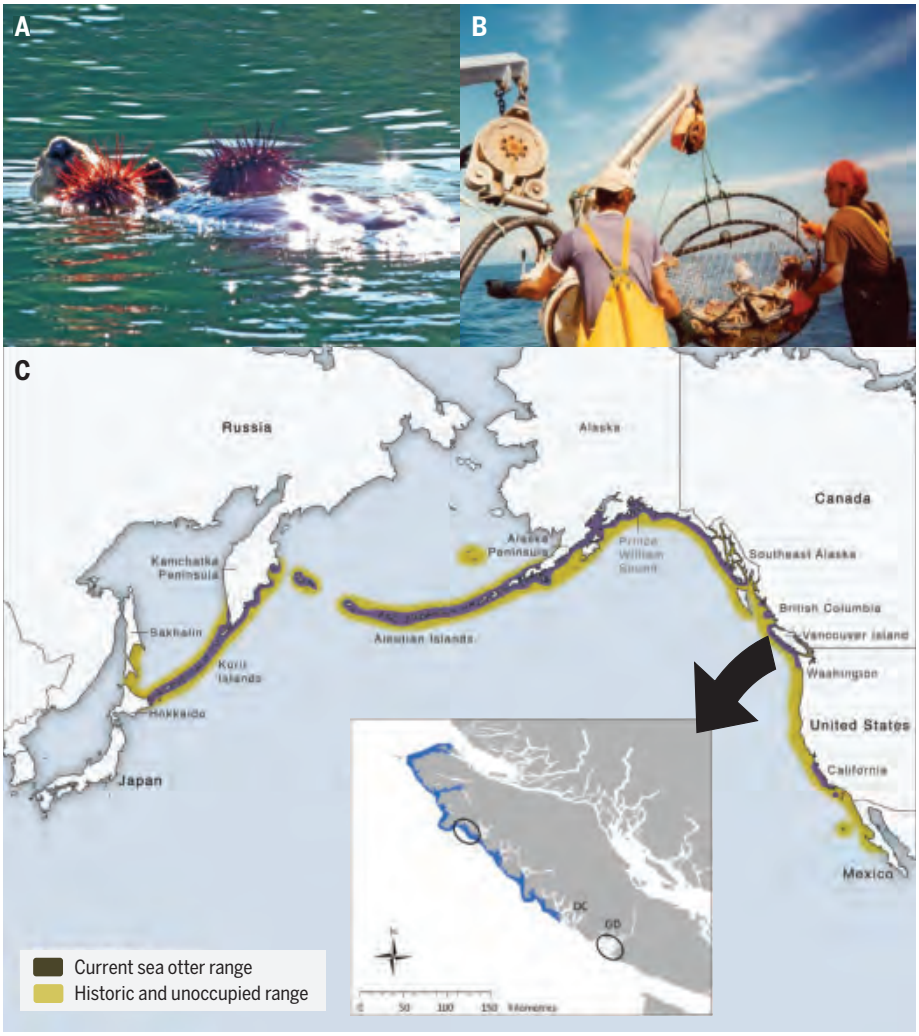


Fig. 1. Ecological and geographic illustration of the study system. (**A**) Sea otter with urchins, a favorite prey item. [Photo credit: B. Nelson] (**B**) Catch of Dungeness crab, a threatened resource. [Photo credit: J. Rogers] (**C**) Range map of historic (yellow) and present-day (dark gray) extents of sea otter distributions in the North Pacific. The inset shows the sea otter range (blue) within the study area, where field data were collected in otter-present and otter-absent areas (ovals), and the location of lucrative Dungeness crab (DC) and geoduck clam (GD) harvesting regions. [Credit: Range map reprinted from (*52*) with permission from Elsevier]

biomass (+37%) reflects the difference between otter-absent and otter-present sites across all groups. Predicted values are reported as median [5th percentile, 95th percentile].

We estimate the lost landed value to commercial invertebrate fisheries from sea otter recovery at 7.3 [4.6, 10.3] million Canadian dollars (CA$)/year (Fig. 3 and table S6). A decline of 25% in the geoduck clam catch comprised over half of this loss. The remainder included the loss of the crab and sea urchin fisheries and a 28% reduction in value to the Manila and butter clam fishery (table S7).

Social and ecological feedbacks (*26*) may mitigate this predicted loss. For example, the global demand for high-value seafood like geoduck clam and Dungeness crab means any reduction in biomass may lead to higher prices, offsetting some of the economic impact to producers. Further, although Dungeness crab largely disappear from our modeled otter-present system, their habitat extends well below the foraging depth of sea otters (*27*). Thus, although lucrative crab fishing grounds in shallow waters will be lost, commercial crab fishers are likely to adapt by shifting fishing efforts to deeper waters.

On the benefits side, costs to the existing fishery are partially offset by a threefold increase in the predicted catch of lingcod, an economically and culturally valuable upper trophic level finfish (Fig. 2 and table S7). More importantly, the increased biomass of kelps and other lower trophic species not explicitly consumed in the model (Fig. 2 and table S8) can yield benefits through deep-ocean carbon storage (*27*) or as a nutritional subsidy to other parts of the ecosystem (*28*, *29*). We estimated the value of this subsidy based on a predicted increase in higher trophic species (i.e., commercial finfish), to be worth 9.4 [2.0, 30.4] million CA$/year (table S6). Uncertainties are high for this service (Fig. 3) because the fate of the surplus production, the trophic transfer efficiencies, and the future landed values are not well known. The estimated value of this service does not include the contribution from increased biomass of subcanopy algal species (*28*), other economic benefits (e.g., recreational fishing, kelp harvesting), or the benefits of the nutritional subsidy to the broader food web.

The portion of unconsumed surplus production lost to deep-ocean storage has value as carbon sequestration. We predict a net benefit of 2.2 [0.5, 7.3] million CA$/year for the sequestered carbon based on European Union carbon prices (Fig. 3 and table S6). This is about one-third of the value obtained by scaling results from a comparable study (*29*) to our study area owing to differences in how kelp production was estimated. Our value can thus be considered a conservative estimate (see supplementary materials for details).

Tourism generates the highest predicted increase in value from sea otter recovery. Our analysis suggests that an otter-dominated system will have the potential to generate a 41.5 [20.7, 66.6] million CA$/year increase in tourism revenue based on willingness-to-pay data derived from a choice experiment (*30*) and recent visitation rates (Fig. 3 and table S6). This estimate does not include likely changes in other tourism-related services such as recreational fishing and destination dive tours. The high uncertainty in this estimate is due to variability in future visitation rates and the estimated willingness to pay. Although this result is based on a local study with existing tourism and sufficient infrastructure to support this increase, other regions in the eastern North Pacific also have established (*12*) or developing (*31*) tourism industries that benefit from the presence of sea otters.

Our estimates of the economic impact of sea otter recovery have wide credibility intervals (Fig. 3), reflecting the uncertainties in parameter values. The distributions of predicted biomass (Fig. 2) were created by randomly resampling the trophic model parameters (see supplementary materials) and show the trophic model was robust to parameter variation. Our social-ecological model combined this uncertainty with other uncertainties, the including the valuation of ecosystem services and trophic interactions among species in the coastal ecosystem. These broad estimates of uncertainty, along with the integration of more generalized models and analyses, combine to improve the representativity of the

results to the broader eastern North Pacific. Although more thorough than many published ecosystem models (*14*), further explorations of model sensitivity to different structures (e.g., trophic networks, valuation methods) would be warranted to support management decisions. Such work must face the challenge of the many poorly understood aspects of social-ecological systems (e.g., nonmonetary values, unknown interactions, nonlinear dynamics, and nonstationarity, including the effects of climate change), which are beyond the scope of the present study.

While acknowledging the limitations of our model, we can be reasonably confident that the otter-present system will yield a higher total economic value, because a net positive outcome is implied across the entire range of the credibility intervals (Fig. 3). This is further supported by empirical evidence showing higher biomass and abundance of many important species in otter-present ecosystems (*9*, *32*–*34*). The uncertainty included in the translation of ecosystem indicators to economic value (see supplementary materials) dominates the uncertainty in the trophic model, as illustrated by the different shapes and credibility intervals for the three services (direct catch, supplemented catch, and carbon; Fig. 3) that depend on the biomass estimates from the trophic model. Our estimates of confidence in the ecological and economic assumptions underlying the service valuations thus provide an intuitive way to visualize the uncertainties associated with such transformations. This approach provides a framework for identifying model components that most limit our understanding of social-ecological systems.

We focused here on the four key monetizable services related to the sea otter trophic cascade. However, such transformations are not valued in a strictly monetary sense by coastal communities (*35*), where social and cultural values are multiple and important (*36*–*38*). Additionally, for coastal communities to benefit from such changes, the resources need to be accessible (*39*, *40*). For example, commercial harvesters generally have the capacity to adapt to shifting resource abundance and distribution, whereas Indigenous or recreational harvesters with more restricted harvesting areas may not be able adapt in the same way. Nor do Indigenous community members necessarily have the ability to access areas (e.g., clam beds) throughout their traditional territories or the capital necessary to take advantage of tourism benefits. Localized losses to subsistence and recreational users can thus be difficult to offset. Given the consolidated nature of invertebrate fisheries in our study area (*41*, *42*) and the relative accessibility of nearshore finfish, the predicted redistribution of biomass from commercial invertebrates to nearshore finfish might be a more equitable distribution of the region's marine productivity. However, the value of tourism, finfish, and invertebrates are not necessarily culturally equivalent to different communities. The benefits of sea otter recovery are therefore likely to be distributed inequitably among economic sectors and local communities, especially of Indigenous peoples, who may experience the losses more acutely than the regional economy as a whole. Although coastal communities in the Pacific Northwest have experienced and adapted to similar shifts in the past (*43*), future adaptation will depend on flexible, multilevel governance structures that allow social-ecological systems to be transformed into more desirable states (*36*, *44*).

Understanding the trade-offs between sea otters and commercial fisheries requires historical context. Today's commercial invertebrate fisheries were made possible by the earlier extirpation of sea otters, which led to hyperabundant populations of large individuals in these target species (*35*), making them an economically viable resource (*5*). The otter-absent system, with its large, abundant invertebrates, thus likely represents a shifted baseline (*45*) for evaluating ecosystem trade-offs, and one that favors the status quo. Nevertheless, the predicted losses to commercial harvesters and coastal communities are legitimate and considerable. Mitigating these social impacts, perhaps by adapting traditional management methods (*36*), could make sea otters less contentious and reduce illegal culling.

Kelp forests also provide additional ecological benefits to the health and productivity of the broader ocean that are outside the scope of our model. Although not yet fully quantified, kelp forests provide habitat to many species and can enhance both biodiversity and resilience (*32*, *46*). The otter-present system would thus seem to support a more resilient social-ecological system, given the increased ecological redundancy and opportunities for diversified fisheries portfolios (*47*).

Further, although our study quantifies the benefits of increased primary production as a nutritional subsidy to one part of the food web (i.e., through catch of valuable finfish), the kelps sustain other coastal species (*48*), as well as pelagic and benthic food webs, because nearly half of the kelp production is estimated to be exported offshore (*49*). How this allochthonous carbon is partitioned between the various food webs and deep-ocean storage remains to be determined. However, it is clear that some coastal regions, including our study area (*50*), export considerable biomass to the open ocean. We therefore propose that kelp-dominated nearshore areas likely serve as primary production pumps and are thus more valuable to

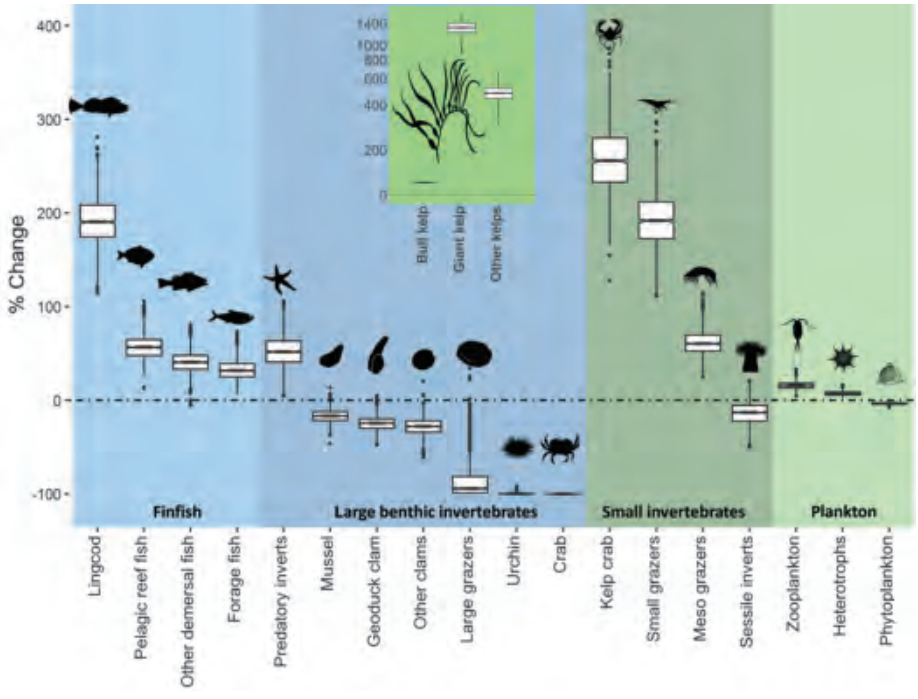


Fig. 2. Percent change in biomass from an otter-absent to an otter-present system. Kelp groups (order Laminariales) are shown as an inset to accommodate the much larger relative biomass change. Functional groups are organized by trophic position (colored background) and then ordered by proportional change, illustrating the switch from a benthic to a more pelagic system and the unaccounted-for surplus biomass in small invertebrates—a key component of the nutritional subsidy to the supplemented catch service (Fig. 3). Boxplots show the range of values resulting from an exploration of valid parameterizations (see supplementary materials for details).

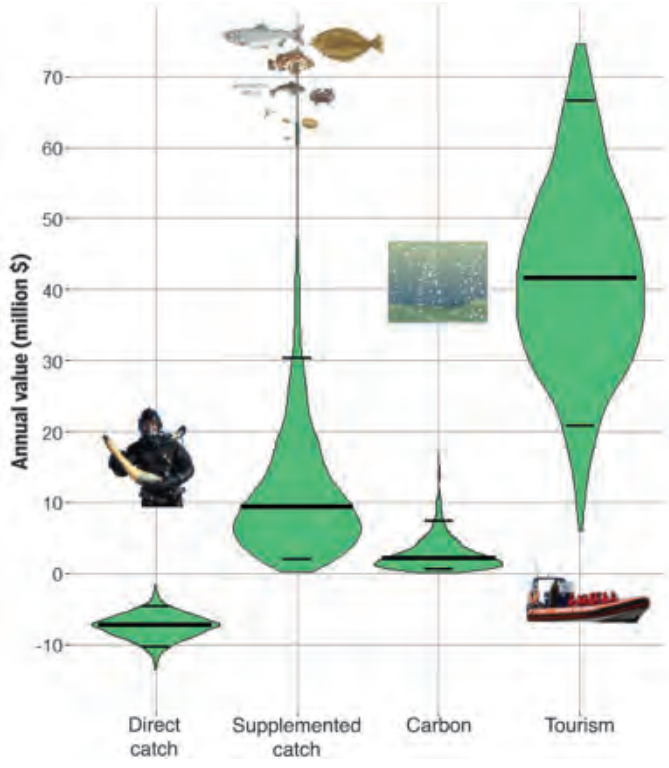


Fig. 3. Sea otter–induced change in annual value for the four ecosystem services considered in this analysis: direct catch, supplemented catch, carbon, and tourism. Changes in value, represented as the difference (in 2018 CA$) between ecosystems with and without sea otters, are shown as violin plots, where the relative widths of the each plot represent the probability distribution of the prediction (like a histogram). The mean and 5th and 95th percentiles are shown as horizontal lines and can be considered the credibility intervals for each service value. These credibility intervals include uncertainties related to the trophic model and in the steps applied to translate the resulting change in ecosystem service supply to dollar values. The intervals reflect the confidence associated with the production and value of each service. Graphical elements illustrate key aspects of each service: Geoduck clams are collected as part of a dive fishery and are the highest value invertebrates in the direct catch; the supplemented catch is defined by a trophic flow to valued finfish such as salmon and halibut (shown at the top of a food chain); marine carbon deposition is principally in the form of marine snow; and wildlife viewing trips are the most conspicuous component of the economic benefits to tourism. [Photo credit: Geoduck diver image provided by the Geoduck Harvesters Association and used with permission]

the world's oceans than previously described [e.g., (*51*)].

The social-ecological model we developed allows the assessment of important social and ecological trade-offs, providing insights into the changes resulting from the recovery of sea otters in the eastern North Pacific. Although the four services we considered (existing invertebrate commercial fisheries, tourism, supplemented finfish catch, and carbon sequestration) do not represent a comprehensive assessment of the social-ecological system, they do provide an innovative perspective on the value of the two ecosystem states. Such integration of diverse services provides a stepping stone toward more complete cost-benefit analyses. Importantly, our broad representation of uncertainty shows how confidence in social-ecological models can be expressed in an intuitive and comprehensible way, allowing meaningful comparisons while illustrating the breadth of uncertainty inherent in such models. Our findings illustrate how sea otters, like many carnivores, exert an oversized effect on social-ecological systems. Hence, coupled social-ecological models are needed for accurately assessing the trade-offs that accompany the loss or recovery of top carnivores in dynamic, continuously adapting systems. Quantifying the impacts of such transformations will inform adaptive management, help mitigate conflicts, promote public acceptance of ecosystem change, and help identify alternate opportunities for local communities.

REFERENCES AND NOTES

1. M. E. Power *et al*., *Bioscience* **46**, 609–620 (1996).
2. J. A. Estes *et al*., *Science* **333**, 301–306 (2011).
3. P. J. Seddon, C. J. Griffiths, P. S. Soorae, D. P. Armstrong, *Science* **345**, 406–412 (2014).
4. A. Treves, K. U. Karanth, *Conserv. Biol.* **17**, 1491–1499 (2003).
5. A. M. Cisneros-Montemayor, M. Barnes-Mauthe, D. Al-Abdulrazzak, E. Navarro-Holm, U. R. Sumaila, *Oryx* **47**, 381–388 (2013).
6. M. Verma *et al*., *Ecosyst. Serv.* **26**, 236–244 (2017).
7. W. J. Ripple, R. L. Beschta, *Biol. Conserv.* **145**, 205–213 (2012).
8. J. A. Estes, J. F. Palmisano, *Science* **185**, 1058–1060 (1974).
9. J. C. Watson, J. A. Estes, *Ecol. Monogr.* **81**, 215–239 (2011).
10. J. C. Watson, T. G. Smith, *Can. Tech. Rep. Fish. Aquat. Sci.* **2089**, 262–303 (1996).
11. A. M. Johnson, in *Transactions of the Forty-Seventh North American Wildlife and Natural Resources Conference*, K. Sabol, Ed. (U.S. Fish and Wildlife Service, 1982), vol. 42, pp. 293–299.
12. J. Loomis, *Coast. Manage.* **34**, 387–404 (2006).
13. F. Turkelboom *et al*., *Ecosyst. Serv.* **29**, 566–578 (2018).
14. E. J. Gregr, K. M. A. Chan, *Bioscience* **65**, 43–54 (2014).
15. K. M. Chan, T. Satterfield, J. Goldstein, *Ecol. Econ.* **74**, 8–18 (2012).
16. K. J. Bagstad, D. J. Semmens, S. Waage, R. Winthrop, *Ecosyst. Serv.* **5**, 27–39 (2013).
17. R. B. Norgaard, *Ecol. Econ.* **69**, 1219–1227 (2010).
18. P. M. Cury, V. Christensen, *ICES J. Mar. Sci.* **62**, 307–310 (2005).
19. E. A. Fulton, A. D. M. Smith, A. E. Punt, *ICES J. Mar. Sci.* **62**, 540–551 (2005).
20. H. Vandermeulen, D. Cobb, *Ocean Coast. Manage.* **47**, 243–256 (2004).
21. S. A. Levin, J. Lubchenco, *Bioscience* **58**, 27–32 (2008).
22. M. A. Zacharias, J. C. Roff, *Conserv. Biol.* **14**, 1327–1334 (2000).
23. I. Rombouts *et al*., *Ecol. Indic.* **24**, 353–365 (2013).
24. M. Coll *et al*., *Ecol. Indic.* **60**, 947–962 (2016).
25. M. Ruckelshaus *et al*., *Ecol. Econ.* **115**, 11–21 (2015).
26. S. Levin *et al*., *Environ. Dev. Econ.* **18**, 111–132 (2012).
27. J. L. Bodkin, G. G. Esslinger, D. H. Monson, *Mar. Mamm. Sci.* **20**, 305–321 (2004).
28. E. U. Rechsteiner, S. B. Wickham, J. C. Watson, *Ecosphere* **9**, e02271 (2018).
29. C. C. Wilmers, J. A. Estes, M. Edwards, K. L. Laidre, B. Konar, *Front. Ecol. Environ.* **10**, 409–415 (2012).
30. R. G. Martone, R. Naidoo, T. Coyle, B. Stelzer, K. M. Chan, *Aquat. Conserv.: Mar. Freshwat. Ecosyst.* (2020).
31. L. Cerveny, *Nature and Tourists in the Last Frontier: Local Encounters with Global Tourism in Coastal Alaska*. (Cognizant Communication Corporation, 2008).
32. R. W. Markel, J. B. Shurin, *Ecology* **96**, 2877–2890 (2015).
33. S. E. Reisewitz, J. A. Estes, C. A. Simenstad, *Oecologia* **146**, 623–631 (2006).
34. J. A. Estes, D. O. Duggins, *Ecol. Monogr.* **65**, 75–100 (1995).
35. M. Fabinyi, W. H. Dressler, M. D. Pido, *Mar. Policy* **94**, 89–92 (2018).
36. J. M. Burt *et al*., *People Nat.* 10.1002/pan3.10090 (2020).
37. M. R. Poe, K. C. Norman, P. S. Levin, *Conserv. Lett.* **7**, 166–175 (2014).
38. S. C. Klain, K. M. A. Chan, *Ecol. Econ.* **82**, 104–113 (2012).
39. N. J. Bennett, H. Govan, T. Satterfield, *Mar. Policy* **57**, 61–68 (2015).
40. R. Wieland, S. Ravensbergen, E. J. Gregr, T. Satterfield, K. M. Chan, *Ecol. Econ.* **121**, 175–180 (2016).
41. A. R. Haas, D. N. Edwards, U. R. Sumaila, *Mar. Policy* **68**, 83–90 (2016).
42. E. Pinkerton, D. N. Edwards, *Mar. Policy* **33**, 707–713 (2009).
43. A. K. Salomon, J. W. Kii'iljuus Barb, X. E. White, N. Tanape Sr, T. M. Happynook, in *Sea Otter Conservation*, S. Larson, J. Bodkin, G. VanBlaricom, Eds. (Elsevier, 2015), pp. 301–331.
44. C. Folke, T. Hahn, P. Olsson, J. Norberg, *Annu. Rev. Environ. Resour.* **30**, 441–473 (2005).
45. D. Pauly, *Trends Ecol. Evol.* **10**, 430 (1995).
46. R. S. Steneck *et al*., *Environ. Conserv.* **29**, 436–459 (2002).
47. T. J. Cline, D. E. Schindler, R. Hilborn, *Nat. Commun.* **8**, 14042 (2017).
48. D. O. Duggins, C. A. Simenstad, J. A. Estes, *Science* **245**, 170–173 (1989).
49. A. Ortega *et al*., *Nat. Geosci.* **12**, 748–754 (2019).
50. D. M. Ware, R. E. Thomson, *Science* **308**, 1280–1284 (2005).
51. T. Agardy *et al*., in *Ecosystems and Human Well-being: Current State and Trends* (Island Press, 2005), vol. 1, pp. 513–549.
52. S. E. Larson, J. L. Bodkin, G. R. VanBlaricom, *Sea Otter Conservation* (Academic Press, 2014).

ACKNOWLEDGMENTS

We thank the communities of Kyuquot and Bamfield, British Columbia, for their hospitality and the Ka:'yu:'k't'h'/Che:k:tles7et'h' and Huu-ay-aht First Nations for providing access to their territories for this work. H. Eyster produced the species outlines in Fig. 2. These are available through Phylopics. N. Bennett provided valuable insights into the nature of nonmonetary benefits, and C. Bullen's careful review improved the quality of this manuscript. We also thank the four anonymous reviewers whose comments improved both the relevance and reproducibility of this work. **Funding:** E.J.G. was funded by a Canadian Graduate Scholarship (DG 06-5566); K.M.A.C., E.J.G., J.B.S., E.A.P.,

C.D.G.H., R.G.M., and R.W.M. were funded by a Canadian Natural Science and Research Council Strategic Grant (SPG 08-1970) to K.M.A.C., C.D.G.H., E.A.P., and J.B.S. We also acknowledge a Canadian NSERC Discovery Grant (RGPIN-2015-05105, K.M.A.C.).
Author contributions: E.J.G., K.M.A.C., C.D.G.H., J.B.S., E.A.P., L.N., R.G.M., R.W.M., and J.C.W. conceived of the study; E.J.G., E.A.P., and V.C. designed the trophic models; R.G.M., R.W.M., J.B.S., E.A.P., C.D.G.H., and E.J.G. collected the data as part of the British Columbia Coastal Ecosystem Services project; J.C.W. and L.N. contributed critical local knowledge and longitudinal data; E.J.G., K.M.A.C., E.A.P., and V.C. conducted the analysis; E.J.G. wrote the manuscript with substantial contributions from all co-authors.
Competing interests: R.G.M. is currently a marine conservation scientist with, and R.W.M. is currently the owner and operator of, Outer Shores Expeditions, a wildlife and cultural expedition tourism company that operates on the British Columbia coast.
Data and materials availability: All data are available in the manuscript or the supplementary materials.

SUPPLEMENTARY MATERIALS
science.sciencemag.org/content/368/6496/1243/suppl/DC1
Materials and Methods
Tables S1 to S8
References (*53*–*111*)
Data S1 and S2

26 June 2019; accepted 5 May 2020
10.1126/science.aay5342

MOLECULAR MACHINES

A precise polyrotaxane synthesizer

Yunyan Qiu[1], Bo Song[1,2], Cristian Pezzato[1]*, Dengke Shen[1], Wenqi Liu[1], Long Zhang[1], Yuanning Feng[1], Qing-Hui Guo[1], Kang Cai[1], Weixingyue Li[1], Hongliang Chen[1], Minh T. Nguyen[1], Yi Shi[1], Chuyang Cheng[1]†, R. Dean Astumian[3]‡, Xiaopeng Li[2]‡, J. Fraser Stoddart[1,4,5]‡

Mechanically interlocked molecules are likely candidates for the design and synthesis of artificial molecular machines. Although polyrotaxanes have already found niche applications in exotic materials with specialized mechanical properties, efficient synthetic protocols to produce them with precise numbers of rings encircling their polymer dumbbells are still lacking. We report the assembly line–like emergence of poly[*n*]rotaxanes with increasingly higher energies by harnessing artificial molecular pumps to deliver rings in pairs by cyclical redox-driven processes. This programmable strategy leads to the precise incorporation of two, four, six, eight, and 10 rings carrying 8+, 16+, 24+, 32+, and 40+ charges, respectively, onto hexacationic polymer dumbbells. This strategy depends precisely on the number of redox cycles applied chemically or electrochemically, in both stepwise and one-pot manners.

The burgeoning of research on mechanically interlocked molecules (MIMs) (*1*) has contributed to progress in the design and synthesis of artificial molecular machines (AMMs) (*2*, *3*). The growing impact of the mechanical bond (*4*) on polymer chemistry and materials science, on the other hand, has led to the development of macroscopic soft materials with unusual properties (*5*). Among these materials, mechanically interlocked polymers (MIPs), such as polyrotaxanes (*6*, *7*) and polycatenanes (*8*, *9*), with complex architectures and topologies, respectively, are desirable synthetic targets that permit access to applications in areas such as slide-ring gels (*10*), battery electrode materials (*11*), and drug delivery platforms (*12*, *13*). These MIPs are usually prepared by template-directed protocols (*14*) that rely on molecular recognition (*15*) and self-assembly (*16*), depending on the synergy between supramolecular chemistry and polymer science.

One of the simplest and most general synthetic strategies for preparing polyrotaxanes is the so-called "threading followed by stoppering" approach (*17*), which takes advantage of noncovalent bonding interactions between polymeric axles and threaded rings with a range of threading kinetics (*18*) to form pseudopolyrotaxanes (*19*), after which bulky stoppers can be connected covalently to both ends of the polymer chains to prevent the loss of the rings by slippage (*20*). The synthesis of polyrotaxanes involving poly(ethylene glycol) (PEG) and cyclodextrins (*17*), exploiting the hydrophobic effect in water, is perhaps the most representative case using this strategy, affording slide-ring materials with remarkable physical properties (*10*) attributable to their mobile ring components. However, certain limitations intrinsically associated with the nature of this synthetic approach appear to be evident, such as (i) restricted control over the number and density of the threaded rings and (ii) limited access to other types of polymer axles displaying weaker or no appreciable noncovalent bonding interactions with the rings.

Oligo- and polyrotaxanes with partial or restricted control over the numbers and densities of the rings have been obtained by several approaches, such as template-directed clipping reactions (*21*, *22*), dynamic acyclic diene metathesis polymerization (*23*), ring-opening olefin metathesis polymerization (*24*), covalent synthesis (*25*), and iterative active-template synthesis (*26*). These strategies can regulate, to differing extents, the numbers, densities, and/or positions of the rings encircling oligo- and polymer axles, leading to enhanced control and distribution of the mobile rings, which serve as a crucial factor in determining and tuning the physical properties of these polyrotaxanes (*27*). Nonetheless, synthetic methodologies, involving precise control over the numbers of rings threaded onto polymers that interact weakly with the rings, are few and far between. We envisioned that a precisely controlled synthetic protocol, harnessing AMMs in delivering rings sequentially onto polymer chains, could give rise to MIPs with flawless control over the numbers of mechanical bonds, the formation of which can be independent of the nature of the chosen polymer dumbbells as a result of the high operational reliability of AMMs. Our synthetic strategy receives its inspiration from the redox-driven artificial molecular pump (AMP), which has been used (*28*–*30*) to recruit cyclobis(paraquat-*p*-phenylene) (**CBPQT^{4+}**) rings from bulk solution onto relatively short oligomeric collecting chains, thus resulting in the production of MIMs away from equilibrium.

Here, we report the precision synthesis (Fig. 1) of enthalpically and entropically demanding multicationic poly[*n*]rotaxanes, in which PEG serves as the centrally located collecting chain in the polymer dumbbells and CBPQT^{4+} as the threaded rings; PEG has been chosen and recognized (figs. S82 to S87) to display no appreciable noncovalent bonding interactions with these rings in acetonitrile. Complete control over the number (up to 10) of rings mechanically interlocked with an individual polymer chain has been realized through the incorporation (Fig. 1A) of two redox-operated AMPs, which are covalently attached to both ends of PEG. These polymer dumbbells can therefore confer on the collecting chains the flawless ability to accommodate multiple ring components sequentially. The pumping of these rings onto the polymer dumbbells is powered precisely by the redox-switching properties of the bipyridinium units associated with both the rings and the AMPs, using an energy-ratchet mechanism (*2*, *31*) reported previously (*28*–*30*). In this manner, two **CBPQT^{4+}** rings can be pumped simultaneously onto the collecting polymer chains after each redox cycle. A defined number of rings (2, 4, 6, 8, and 10) can then be installed (Fig. 1B) onto the collecting PEG chains, depending solely on the number of redox cycles applied to both the AMP-functionalized polymer dumbbells and the **CBPQT^{4+}** rings.

The polymer dumbbell **PolyDB^{6+}** (Fig. 2A) incorporating two AMPs at both ends of the azide-terminated PEG (number-average

[1]Department of Chemistry, Northwestern University, Evanston, IL 60208, USA. [2]Department of Chemistry, University of South Florida, Tampa, FL 33620, USA. [3]Department of Physics, University of Maine, Orono, ME 04469, USA. [4]Institute for Molecular Design and Synthesis, Tianjin University, Tianjin 300072, P. R. China. [5]School of Chemistry, University of New South Wales, Sydney, NSW 2052, Australia.
*Present address: Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland.
†Present address: College of Chemistry, Sichuan University, Chengdu, Sichuan 610064, P. R. China.
‡Corresponding author. Email: astumian@maine.edu (R.D.A.); xiaopengli1@usf.edu (X.L.); stoddart@northwestern.edu (J.F.S.)

molecular weight M_n = 2000, polydispersity $Đ$ = 1.05) (*32*) was synthesized (fig. S1) using click chemistry (*33*) and has been characterized (fig. S3) by ^{1}H nuclear magnetic resonance (NMR) spectroscopy. The AMP (Fig. 2A) consists of a bipyridinium ($BIPY^{2+}$) unit positioned between a 2,6-dimethylpyridinium (PY^+) coulombic barrier and an isopropylphenylene (IPP) steric barrier, which, taken as a trio, gives access to the formation of mechanical bonds with the $CBPQT^{4+}$ rings. The operation of the two-terminal AMPs relies on the redox-switching properties of $BIPY^{2+}$ units associated with both the molecular pumps and the rings. Upon reduction, all $BIPY^{2+}$ units are reduced to their radical cationic states, leading to the threading of two **CBPQT$^{2(+\bullet)}$** rings onto **PolyDB$^{4(+)2(\bullet)}$** (Fig. 2B, II), with the formation of two trisradical tricationic complexes based on radical-radical interactions (*34*). After oxidation, the strong coulombic repulsions between charged units force the rings to traverse the IPP units (Fig. 2B, III) by exploiting the imbalance between the energy barriers exerted from the charged (PY^+ and $BIPY^{2+}$) units of the pumps and the neutral IPP units, resulting in the formation of the poly[3]rotaxane during thermal relaxation (Fig. 2B, IV). A second redox cycle following an analogous mechanism produces the poly[5]rotaxane (Fig. 2B, V, VI, and VII), and so on.

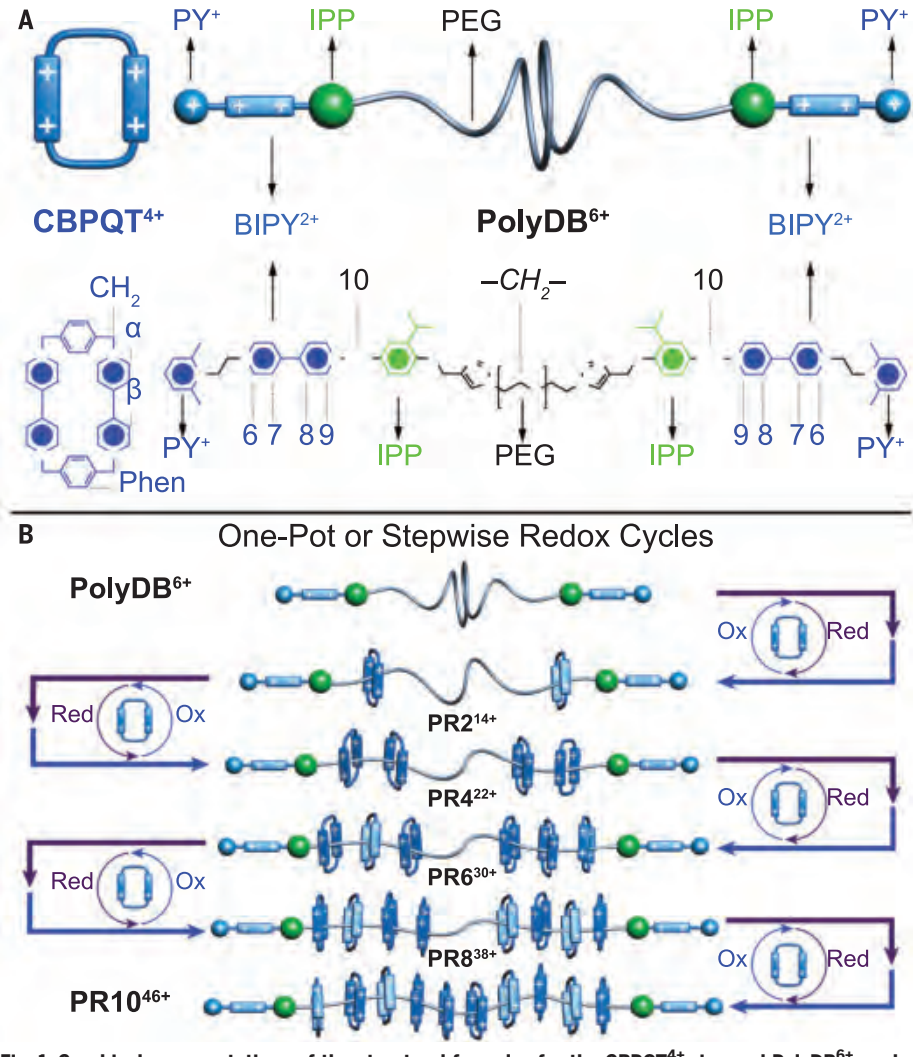


Fig. 1. Graphical representations of the structural formulas for the CBPQT^{4+} ring and PolyDB^{6+}, and the redox-driven syntheses of poly[*n*]rotaxanes using AMPs. (**A**) Schematic representations of the structural formulas for the **CBPQT^{4+}** ring and **PolyDB^{6+}**, which consists of two molecular pumps attached at both ends of the PEG chains through a pair of triazole rings, forming a collecting polymer chain located in the middle between the two pumps. The PF_6^- counterions are omitted for the sake of clarity. Selected key protons on **CBPQT^{4+}** (Phen, CH_2, α, and β) and **PolyDB^{6+}** (numbers 6 to 10) are labeled to aid the interpretation of ^{1}H NMR spectra reproduced in Fig. 3. The assignments of all proton resonances of **PolyDB^{6+}** are labeled in fig. S3. (**B**) Schematic representation of the repetitive reduction and oxidation processes, either stepwise or in a one-pot synthesis, producing a series of polyrotaxanes **PR2^{14+}**, **PR4^{22+}**, **PR6^{30+}**, **PR8^{38+}**, and **PR10^{46+}** with 2, 4, 6, 8, and 10 rings, respectively.

The polyrotaxane synthesizer is operated by an energy-ratchet mechanism (*2*, *31*). The fundamental operation principle relies on the fact that the energy of the electrostatic barriers arising from the PY^+ and $BIPY^{2+}$ units and of the radical-pairing interactions forming trisradical tricationic complexes depends strongly on the redox state of the system, whereas the heights of the steric barriers imposed by the IPP units are independent of the redox state (Fig. 2B). The resulting kinetic asymmetry (fig. S2) allows the pumps to use the energy supplied by the alternating redox reagents or potentials to drive the formation of a highly nonequilibrium mechanically interlocked system—a polyrotaxane with up to 10 rings threaded onto a collecting chain—that persists as a metastable state away from equilibrium for a long time.

First, we subjected **PolyDB^{6+}** and a large excess of **CBPQT^{4+}** to one cycle of redox reactions using cobaltocene (reductant) and $NOPF_6$ (oxidant) (*35*). Subsequent ^{1}H NMR spectroscopic analysis (fig. S4) confirmed the successful production of the poly[3]rotaxane **S-PR2^{14+}** (**S** denotes stepwise) with only two mechanically interlocked $CBPQT^{4+}$ rings, showing that both AMPs operate simultaneously. Motivated by this accomplishment, we decided to apply a stepwise synthetic strategy to the isolated poly[3]rotaxane **S-PR2^{14+}** that was used as the starting material, and subsequently generated the poly[5]rotaxane **S-PR4^{22+}** with four $CBPQT^{4+}$ rings after a second redox cycle. A repeat of this redox cycle led to the poly[7]rotaxane **S-PR6^{30+}** with six mechanically interlocked rings, using **S-PR4^{22+}** as the starting material. All three poly[*n*]rotaxanes have been characterized (figs. S4 to S6) by ^{1}H NMR spectroscopy.

The ease of the redox operation of the AMPs simply by the addition of redox reagents, as well as a relatively short operating time (within 1 hour) for each redox cycle, also led us to the adoption of a one-pot synthetic strategy, which eliminates the tedious isolation of intermediate poly[*n*]rotaxane precursors. Accordingly, **O-PR2^{14+}**, **O-PR4^{22+}**, and **O-PR6^{30+}** (**O** denotes one-pot) can be produced by subjecting **PolyDB^{6+}** to one, two, and three redox cycles, respectively, through the repetitive, alternating addition of reductant and oxidant into the reaction mixture. We then continued the redox cycle, producing the poly[9]rotaxane **O-PR8^{38+}** and the poly[11]rotaxane **O-PR10^{46+}** with corresponding numbers of 8 and 10 rings, respectively. All these poly[*n*]rotaxanes produced by one-pot redox cycles have been characterized quantitatively (Fig. 3 and figs. S7 to S16 and S23 to S29) using ^{1}H NMR spectroscopy and ^{1}H-^{1}H correlation spectroscopy (COSY) with mesitylene as an internal standard. Integration of the resulting ^{1}H NMR spectra allowed us to calculate the M_n values of the homologous series (Table 1) of these poly[*n*]rotaxanes.

These calculated results are in good agreement with the theoretical M_n.

We also used 2D diffusion-ordered spectroscopy (DOSY) (figs. S30 to S35) and nuclear Overhauser effect spectroscopy (NOESY) measurements (figs. S36 to S40) to confirm the mechanically interlocked nature of the poly[n]rotaxanes. It is clear from Fig. 3, A to E, that the resonances for the $-OCH_2CH_2-$ repeating units in the PEG backbones undergo a continuous upfield shift with the increasing number of rings, as a result of the amplification of the shielding effect exerted by the $CBPQT^{4+}$ rings. The number of threaded rings can be estimated from the integration of probe resonances in the ^1H NMR spectra recorded in Fig. 3. The ratio of the rings to the polymer dumbbell is deduced from carrying out comparisons between the expected integration values (numbers in parentheses in Fig. 3) and the actual values (numbers under peaks in Fig. 3) of proton resonances present in both the mechanically interlocked $CBPQT^{4+}$ rings ($H_{\alpha'}$, $H_{\beta'}$, $H_{Phen'}$, and $H_{CH2'}$) and the $PolyDB^{6+}$ components ($H_{6'}$, $H_{9'}$, $H_{7'}$, $H_{8'}$, and $H_{10'}$). In all five cases, the actual integrated values of the proton resonances are in good agreement with the expected ones, thus indicating the successful threading of the rings onto $PolyDB^{6+}$ with the targeted number (2, 4, 6, 8, and 10). Among these ^1H NMR spectra, those (Fig. 3, D and E) of **O-PR8**$^{38+}$ and **O-PR10**$^{46+}$ exhibit signal broadening attributable to the increased number of threaded rings, resulting in small deviations (figs. S11 and S14) between the expected and actual integrated values associated with

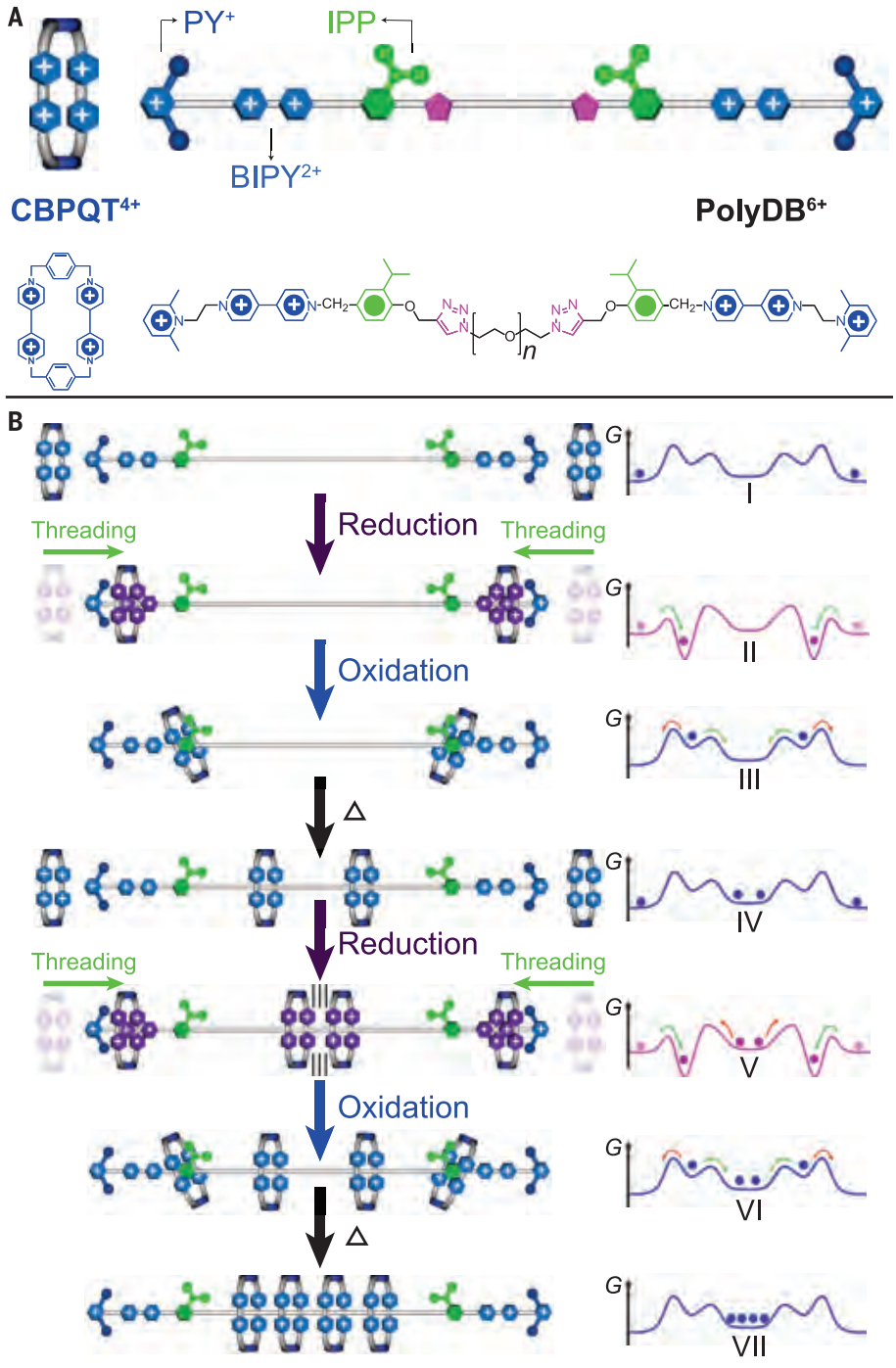


Fig. 2. Graphical representations of the redox-driven synthesis of the poly[5]rotaxane by means of a flashing energy-ratchet mechanism. (**A**) Graphical representations of the structural formulas for the **CBPQT**$^{4+}$ ring and **PolyDB**$^{6+}$, which consists of two molecular pumps attached at both ends of the PEG chains by a pair of triazole rings, forming a collecting polymer chain located in the middle between the two pumps. The PF_6^- counterions are omitted for the sake of clarity. (**B**) The pumping mechanism for the redox-driven synthesis of the poly[5]rotaxane. Energy profiles representing the free energies of the system, as the rings are pumped in pairs onto the polymer dumbbell, are illustrated to the right of each intermediate in the reaction sequence. The curved arrows on the energy profiles represent reaction pathways that are either kinetically favored (green) or disfavored (red). The triazole rings (magenta) in each intermediate of the reaction are omitted for the sake of clarity. Initially (I), the rings and the polymer dumbbell repel each other because of strong coulombic repulsions. Upon reduction (II), all $BIPY^{2+}$ units are reduced to their radical cationic states ($BIPY^{+\bullet}$), leading to the threading of two **CBPQT**$^{2(+\bullet)}$ rings, one onto each end of **PolyDB**$^{4(+)2(\cdot)}$ with the formation of two trisradical tricationic complexes. Upon oxidation (III), the strong coulombic repulsions between the charged PY^+, $BIPY^{2+}$, and $CBPQT^{4+}$ units lead to the rings traversing the IPP steric barriers as a result of thermal energy, and falling into a kinetic trap provided by the collecting polymer chain, resulting in the formation (IV) of the poly[3]rotaxane. The second reduction (V) allows the pumps to recruit two more rings from the bulk solution by a similar mechanism. (In addition, the two threaded reduced rings most likely interact through radical-pairing interactions, indicated by a pair of triple vertical lines.) The second oxidation (VI) restores full charges to the $BIPY^{2+}$ units and the $CBPQT^{4+}$ rings and forces the second pair of rings to traverse the IPP steric barriers while the two reoxidized threaded rings are prevented from dethreading by these steric barriers, aided by the strong coulombic repulsions associated with both pumps. The second thermal relaxation results in the formation (VII) of the poly[5]rotaxane, and so on.

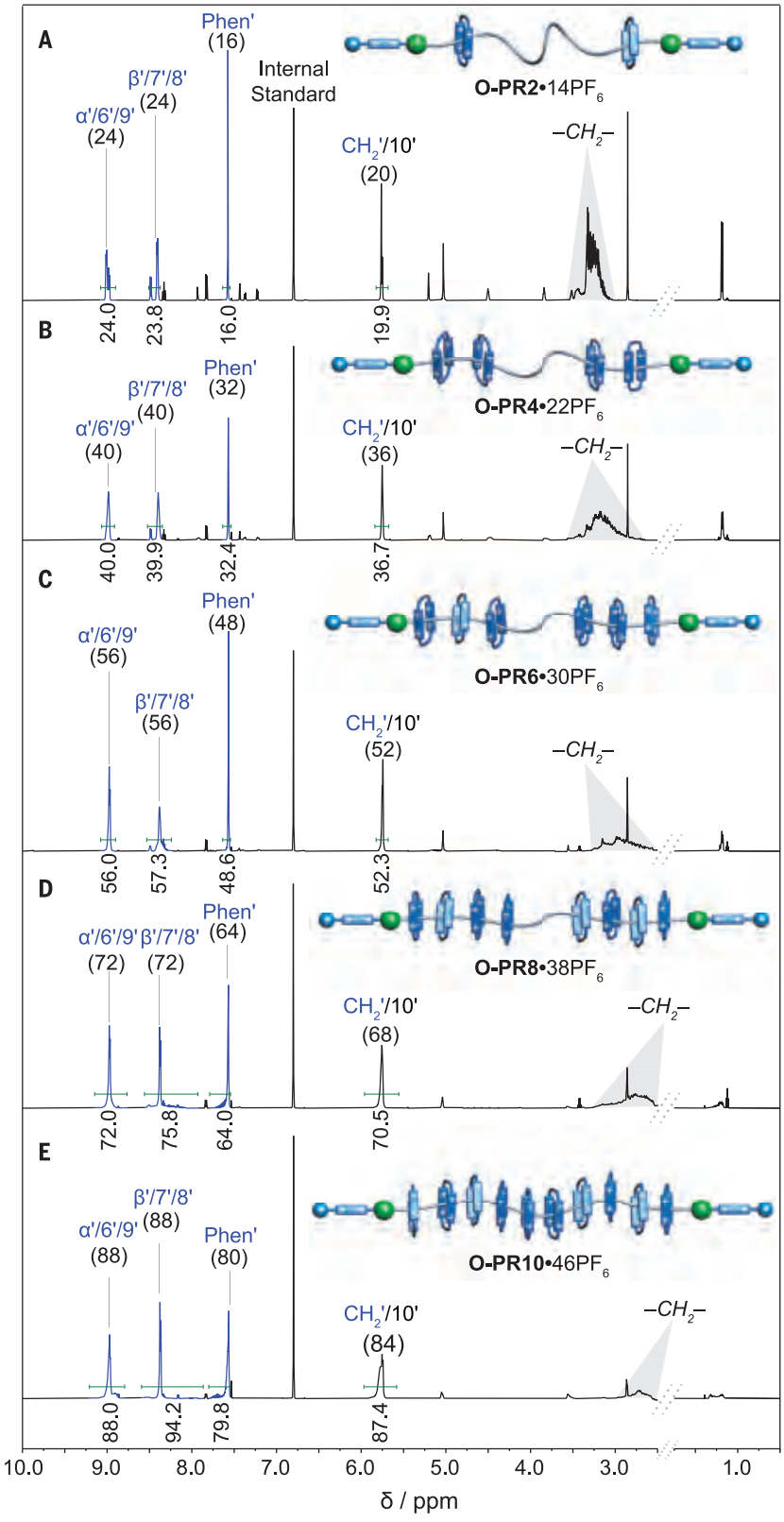


Fig. 3. NMR spectroscopic characterization of all five poly[*n*]rotaxanes. (**A**) Partial ^{1}H NMR spectrum (600 MHz, CD_3CN, 298 K) of **O-PR2**$^{14+}$. (**B**) Partial ^{1}H NMR spectrum (600 MHz, CD_3CN, 298 K) of **O-PR4**$^{22+}$. (**C**) Partial ^{1}H NMR spectrum (600 MHz, CD_3CN, 298 K) of **O-PR6**$^{30+}$. (**D**) Partial ^{1}H NMR spectrum (600 MHz, CD_3CN, 298 K) of **O-PR8**$^{38+}$. (**E**) Partial ^{1}H NMR spectrum (600 MHz, CD_3CN, 298 K) of **O-PR10**$^{46+}$. Numbers in parentheses represent the expected integration values of the proton resonances of interest; numbers under the peaks represent the actual integrated values. Primes denote the proton resonances of mechanically interlocked molecules.

the proton resonances for $H_{\beta'}$, $H_{7'}$, and $H_{8'}$. These deviations were mitigated after deconvolution to remove the integrations from irrelevant overlapping resonances (figs. S12 and S15), revealing excellent agreement between the actual integrated values and the expected ones.

To obtain unambiguous evidence of the coconstitutions of these poly[*n*]rotaxanes and determine the value of *n* quantitatively, we used electrospray ionization–mass spectrometry (ESI-MS) of the polymer dumbbell precursor—the PEG bis(azide) (M_n = 2000)—to probe the mass-to-charge ratio (*m*/*z*) distribution as a function of the different numbers of repeating $–OCH_2CH_2–$ units in each homogeneous collection of polymer chains. Three sets of signals, which are independently distributed without overlapping with each other, are observed (fig. S41) for the charge states ranging from 1+ to 3+, corresponding to $[M+Na]^+$, $[M+2Na]^{2+}$, and $[M+2H+Na]^{3+}$, respectively. This information lays the foundation for an in-depth analysis of **PolyDB**$^{6+}$ and the derived poly[*n*]rotaxanes. The ESI-MS analysis of **PolyDB**$^{6+}$ (Fig. 4A, purple) shows primarily four sets of signals ranging from 3+ to 6+, with each peak matching the chemical composition of $[(C_{66}H_{76}N_{12}O_2)(C_2H_4O)_n(PF_6^-)_x]^{(6-x)+}$ (fig. S42 and tables S1 to S4). Another set of signals with the charge state 4+ is also observed (Fig. 4A, green) as a result of the fragmentation of **PolyDB**$^{6+}$. We reasoned that under the mass spectrometric conditions, the $BIPY^{2+}$ units become labile and fragment from one or both ends, generating polymer dumbbells terminated by benzylic cations and the cleaved monobipyridinium fragments (fig. S43). To verify this hypothesis, we applied harsher ESI conditions by increasing both the sampling and extraction cone voltages. A singly charged peak with *m*/*z* = 436.14 appears (fig. S44) as expected, with the isotope pattern matching the theoretical simulation of a dissociated monobipyridinium unit. The ESI-MS analysis of **O-PR2**$^{14+}$ shows the existence of three species (Fig. 4B, fig. S50, and tables S5 to 10). The 5+ and 6+ peaks (Fig. 4B, purple) correspond to the intact poly[3]rotaxane, whereas the green and blue signals are consequences of the fragmentations at one and both ends, respectively. The dissociation of the $BIPY^{2+}$ units from the polymer dumbbells does not jeopardize the mechanically interlocked nature of the poly[3]rotaxane because of the steric hindrance provided by the IPP units. Taken overall, these results, together with the ^{1}H NMR spectroscopic analysis (Fig. 3A), indicate clearly that only two rings are pumped onto each individual polymer dumbbell after one redox cycle.

Analyses of **O-PR4**$^{22+}$, **O-PR6**$^{30+}$, and **O-PR8**$^{38+}$ (Fig. 4, C to E) also reveal mass spectra with primarily two sets of signals corresponding

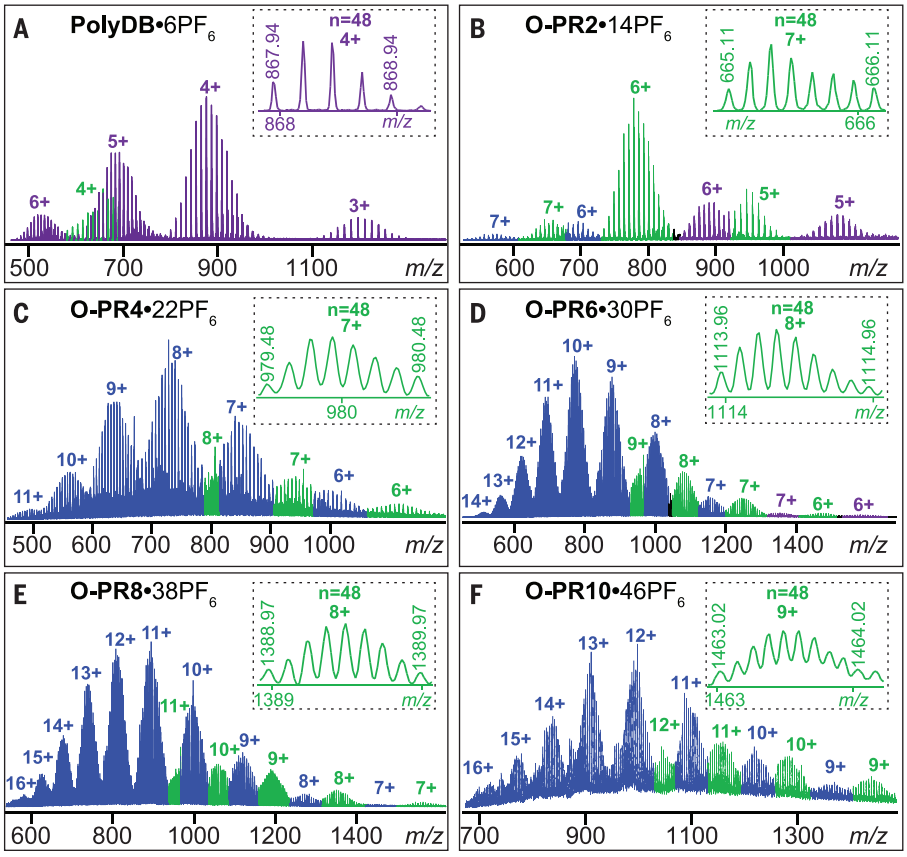


Fig. 4. Electrospray ionization-mass spectrometry (ESI-MS) measurements of PolyDB•6PF_6 and all five poly[*n*]rotaxanes. (**A**) ESI-MS for **PolyDB**•6PF_6. The inset shows the isotope pattern for [(**PolyDB**)•2PF_6]$^{4+}$ with 48 repeating –OCH_2CH_2– units. (**B** to **F**) ESI-MS full spectra for **O-PR2**•14PF_6 (B), **O-PR4**•22PF_6 (C), **O-PR6**•30PF_6 (D), **O-PR8**•38PF_6 (E), and **O-PR10**•46PF_6 (F). The insets show the isotope pattern for [(**O-PR2**)*•7PF_6]$^{7+}$, [(**O-PR4**)*•15PF_6]$^{7+}$, [(**O-PR6**)*•22PF_6]$^{8+}$, [(**O-PR8**)*•30PF_6]$^{8+}$, and [(**O-PR10**)*•37PF_6]$^{9+}$, respectively, with their corresponding charge states. The asterisk represents fragmentation at one end. The differences of these charge states in each spectrum result from the losses of different numbers of counterions (PF_6^-), and the distributions in charge states derive from PEG with different numbers of repeating units. Fragmentations at one or two ends are colored green and blue, respectively. The intact poly[*n*]rotaxanes without fragmentations are colored purple.

to the same fragmentations. Even though the peaks for intact poly[*n*]rotaxanes are hardly observable, possibly as a result of the diminished stability of the $BIPY^{2+}$ units in the case of the larger poly[*n*]rotaxanes, the mechanically interlocked characteristics of these poly[*n*]rotaxanes remain intact. Each charge state shows a clear mass distribution assignable to different numbers of repeating –OCH_2CH_2– units (figs. S51 to S56 and tables S11 to S24). The peaks observed in the spectra can only be assigned to **O-PR4**$^{22+}$, **O-PR6**$^{30+}$, and **O-PR8**$^{38+}$ with the corresponding four, six, and eight rings (*36*) mechanically interlocked with each individual polymer dumbbell, respectively.

O-PR10^{46+}, which has the largest molecular weight among the as-synthesized poly[*n*]rotaxanes, was also analyzed using ESI-MS, revealing a spectrum with only two series of charge states (Fig. 4F). Despite fragmentation, each peak, after deconvolution, matches the chemical composition of the fragmented **O-PR10**$^{46+}$, with the number of repeating –OCH_2CH_2– units representing the only difference (figs. S57 and S58 and tables S25 to S28). It is also evident from careful comparison (figs. S59 and S60) between the two ESI-MS spectra (Fig. 4, E and F) that a trace amount of **O-PR8**$^{38+}$ is present in the sample of **O-PR10**$^{46+}$. This observation implies that the polyrotaxanes with shorter collecting chains could already have reached the steady state in which two threaded $CBPQT^{2(+\bullet)}$ rings on the collecting chain move to encircle the $BIPY^{+\bullet}$ units under reducing conditions and then move back to the collecting chain upon oxidation. Nonetheless, the findings from ESI-MS lend strong support to the successful one-pot precision synthesis of **O-PR10**$^{46+}$ with a well-defined number (10) of rings, affording out-of-equilibrium polycationic poly[11]rotaxane with high charge densities (46+) on single polymer chains.

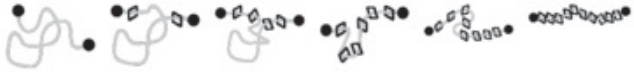
With both stepwise and one-pot chemical redox synthetic strategies established, we turned our attention to developing an additional protocol that could potentially mitigate the accumulation of waste products resulting from the repetitive additions of redox reagents. Recently, we demonstrated (*37*) the use of electricity to operate AMPs in controlled potential electrolysis. Cyclic voltammograms (CVs) of all the poly[*n*]rotaxanes, which revealed nearly identical first reduction potentials, prove (fig. S68) that a similar standard electrochemical protocol using a reduction potential of –0.7 V during electrolysis should be applicable to the production of higher-order poly[*n*]rotaxanes. Ultimately, application of the electrochemical synthetic protocol in a stepwise and controlled manner leads to the production of the poly[*n*]rotaxanes **E-PR2**$^{14+}$, **E-PR4**$^{22+}$, **E-PR6**$^{30+}$, and **E-PR8**$^{38+}$ [(*38*); **E** denotes electrochemical] that are identical in comparison with the chemically synthesized poly[*n*]rotaxanes **O-PR2**$^{14+}$, **O-PR4**$^{22+}$, **O-PR6**$^{30+}$, and **O-PR8**$^{38+}$, respectively. All these electrochemically produced poly[*n*]rotaxanes were fully characterized by ^{1}H NMR spectroscopy (figs. S17 to S21) and ESI-MS (figs. S61 to S67), including confirmation of the targeted number of rings, thereby highlighting the robustness and reliability of AMP-enabled precision synthesis of poly[*n*]rotaxanes by electrochemical means.

Finally, as the number of rings mechanically interlocked along the polymer dumbbells increases with the number of redox cycles, the random-coil backbones of PEG could become progressively more rigid and manifest an increase in their hydrodynamic diameters (D_h). Accordingly, we performed additional experiments, including DOSY, size exclusion chromatography (SEC), and dynamic light scattering (DLS) to probe the differences in size of **PolyDB**$^{6+}$ and the derived poly[*n*]rotaxanes. As anticipated, the 2D DOSY experiments (Table 1 and figs. S30 to S35) revealed a steady decrease in the diffusion constant, from 4.11 × 10^{-6} cm^2/s for **PolyDB**$^{6+}$ to 2.30 × 10^{-6} cm^2/s for **O-PR10**$^{46+}$, indicating an increase in D_h. The SEC elution profiles (figs. S75 to S81) show a decrease in elution time with increasing number of rings. The DLS data (Table 1 and figs. S69 to S74 and S81) confirm the increase in D_h, ranging from 2.67 ± 1.02 nm for **PolyDB**$^{6+}$ to 4.75 ± 1.38 nm for **O-PR10**$^{46+}$. We also carried out zeta potential measurements (Table 1), which show an increasing value from 0.77 ± 1.16 mV for **PolyDB**$^{6+}$ to 14.47 ± 0.35 mV for **O-PR10**$^{46+}$ as the charges accumulate rapidly with the increasing number of threaded rings. These observations, taken as a whole, confirm the progressive increase in hydrodynamic diameters from **PolyDB**$^{6+}$ to **O-PR10**$^{46+}$ as a direct result of the accumulation of the threaded rings with a number precisely defined.

The sequential installation of up to 10 rings onto the polymer dumbbell **PolyDB^{6+}** is an example of redox-driven self-assembly (*39*) (Fig. 1B). From a thermodynamic perspective, the most probable structure arising from five cycles of pumping is **PR10^{46+}**, a highly unstable and energetic far-from-equilibrium polyrotaxane. Without pumping, the most probable structure is the polymer dumbbell **PolyDB^{6+}** because of the decrease in entropy upon threading a pair of rings onto the collecting polymer chain. The redox-driven pumping process, by alternating reduction and oxidation consecutively, inputs energy into the system, which drives the resulting polyrotaxanes away from equilibrium. After oxidation, a pair of rings undergoes thermally activated relaxation, either back into the bulk solution over the PY^+ units (very unlikely) or onto the collecting chain over the IPP units (extremely likely) as a result of the large free energy for rings encircling the $BIPY^{2+}$ units after oxidation. A steady-state regime will be eventually reached after many cycles of redox reactions, where the likelihood of a ring escaping to the bulk solution from the polymer dumbbell is equal to the likelihood of a ring being pumped from the bulk solution onto the polymer dumbbell—the process is a dynamic one—and the most likely outcome during a further redox cycle at steady state is that in which a pair of threaded $CBPQT^{2(+\bullet)}$ rings from the collecting chain moves to encircle the $BIPY^{+\bullet}$ units under reducing conditions and then simply moves back again onto the collecting chain upon oxidation. Although it is tempting to describe this process as "dissipative assembly," this description would be misleading. The key element that allows pumping to maintain a highly nonequilibrium structure is kinetic asymmetry (*40*) provided by the "pumping cassette" (*41*–*43*), which comprises a switchable barrier (PY^+), a switchable recognition site ($BIPY^{2+}/BIPY^{+\bullet}$), and a fixed neutral steric barrier (IPP). If the PY^+ units were substituted by the IPP units or vice versa, redox cycling would still dissipate energy but there would be no pumping of the rings onto the polymer dumbbell. The most probable compound in the presence of redox cycling, without kinetic asymmetry, would be **PolyDB^{6+}** with no threaded rings. (See the supplementary materials for a more detailed discussion.)

The six positive charges of **PolyDB^{6+}**, already associated with the operational components of the pumps on the polymer dumbbells, can be augmented by another 8+ to 16+ to 24+ to 32+ to 40+ with the pairwise addition of two, four, six, eight, and 10 rings, respectively. Hence, emergent behavior on the part of the polyrotaxanes becomes a real possibility. This foray into ever-increasing nonequilibrium states raises unlimited questions and opens doors to a whole range of new directions that can be pursued in unnatural polymer synthesis. There is the opportunity to tune the materials properties of slide-ring polymers with more or fewer rings located at will, and almost free of noncovalent bonding interactions with the polymer dumbbells. The fact that palindromic arrays of co-constitutionally heterotopic rings can be positioned on constitutionally symmetrical dumbbells means that it is now possible to use polyrotaxanes, engineered in an appropriate manner, to transcribe their programmed information back into the domain of controlling, by the use of further templation, sequences of building blocks in a new line of wholly synthetic polymers. These possibilities represent little more than the tip of the iceberg.

Table 1. Molar mass and physical constants of the polymer dumbbell and the derived poly[*n*] rotaxanes. Theoretical and ^{1}H NMR–derived number-average molecular weights (M_n), diffusion constants (D), hydrodynamic diameters (D_h), and zeta potentials (ζ) are shown for **PolyDB^{6+}**, **O-PR2^{14+}**, **O-PR4^{22+}**, **O-PR6^{30+}**, **O-PR8^{38+}**, and **O-PR10^{46+}**. Errors are reported as 1 SD. The theoretical M_n of each sample where the counterions are PF_6^- was estimated on the basis of the assumption that the molecular weight of the collecting polymer chain is 2000 g/mol. ^{1}H NMR and DOSY spectra were measured in CD_3CN on samples where the counterions were PF_6^-, whereas dynamic light scattering and zeta potential measurements were conducted in phosphate-buffered saline with either TFA^- or Cl^- as the counterions.

Sample	PolyDB^{6+}	O-PR2^{14+}	O-PR4^{22+}	O-PR6^{30+}	O-PR8^{38+}	O-PR10^{46+}
Representation						
M_n (theory) (g/mol)	3800	6000	8200	10,400	12,600	14,800
M_n (NMR) (g/mol)	4800	6200	9400	11,800	14,000	16,700
D (cm^2/s)	4.11×10^{-6}	3.31×10^{-6}	2.86×10^{-6}	2.66×10^{-6}	2.48×10^{-6}	2.30×10^{-6}
D_h (nm)	2.67 ± 1.02	2.74 ± 0.90	2.93 ± 1.09	4.02 ± 1.12	4.48 ± 0.74	4.75 ± 1.38
ζ (mV)	0.77 ± 1.16	3.14 ± 0.67	5.03 ± 0.58	6.50 ± 2.34	8.48 ± 0.44	14.47 ± 0.35

REFERENCES AND NOTES

1. J. F. Stoddart, *Angew. Chem. Int. Ed.* **56**, 11094–11125 (2017).
2. C. Pezzato, C. Cheng, J. F. Stoddart, R. D. Astumian, *Chem. Soc. Rev.* **46**, 5491–5507 (2017).
3. S. Erbas-Cakmak, D. A. Leigh, C. T. McTernan, A. L. Nussbaumer, *Chem. Rev.* **115**, 10081–10206 (2015).
4. C. J. Bruns, J. F. Stoddart, *The Nature of the Mechanical Bond: From Molecules to Machines* (Wiley, 2016).
5. S. Mena-Hernando, E. M. Pérez, *Chem. Soc. Rev.* **48**, 5016–5032 (2019).
6. A. Harada, A. Hashidzume, H. Yamaguchi, Y. Takashima, *Chem. Rev.* **109**, 5974–6023 (2009).
7. M. Arunachalam, H. W. Gibson, *Prog. Polym. Sci.* **39**, 1043–1073 (2014).
8. Z. Niu, H. W. Gibson, *Chem. Rev.* **109**, 6024–6046 (2009).
9. Q. Wu *et al.*, *Science* **358**, 1434–1439 (2017).
10. Y. Noda, Y. Hayashi, K. Ito, *J. Appl. Polym. Sci.* **131**, 40509 (2014).
11. S. Choi, T.-w. Kwon, A. Coskun, J. W. Choi, *Science* **357**, 279–283 (2017).
12. K. K. Cotí *et al.*, *Nanoscale* **1**, 16–39 (2009).
13. J. J. Li, F. Zhao, J. Li, *Appl. Microbiol. Biotechnol.* **90** 427–443 (2011).
14. F. M. Raymo, J. F. Stoddart, *Chem. Rev.* **99**, 1643–1664 (1999).
15. D. J. Cram, *Science* **240**, 760–767 (1988).
16. J.-M. Lehn, *Science* **295**, 2400–2403 (2002).
17. A. Harada, J. Li, M. Kamachi, *Nature* **356**, 325–327 (1992).
18. A. B. C. Deutman *et al.*, *Science* **322**, 1668–1671 (2008).
19. F. M. Raymo, J. F. Stoddart, *Trends Polym. Sci.* **4**, 208–211 (1996).
20. P. R. Ashton, M. Bělohradský, D. Philp, J. F. Stoddart, *J. Chem. Soc. Chem. Commun.* **1993**, 1269–1274 (1993).
21. J. Wu, K. C.-F. Leung, J. F. Stoddart, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 17266–17271 (2007).
22. M. E. Belowich *et al.*, *J. Am. Chem. Soc.* **134**, 5243–5261 (2012).
23. N. Momčilović, P. G. Clark, A. J. Boydston, R. H. Grubbs, *J. Am. Chem. Soc.* **133**, 19087–19089 (2011).
24. S. Kang, M. M. Cetin, R. Jiang, E. S. Clevenger, M. F. Mayer, *J. Am. Chem. Soc.* **136**, 12588–12591 (2014).
25. M. D. Cornelissen *et al.*, *J. Org. Chem.* **85**, 3146–3159 (2020).
26. J. E. M. Lewis, J. Winn, L. Cera, S. M. Goldup, *J. Am. Chem. Soc.* **138**, 16329–16336 (2016).
27. K. Mayumi, K. Ito, K. Kato, *Polyrotaxane and Slide-Ring Materials* (RSC, 2016).
28. C. Cheng *et al.*, *Nat. Nanotechnol.* **10**, 547–553 (2015).
29. C. Pezzato *et al.*, *Tetrahedron* **73**, 4849–4857 (2017).
30. Y. Qiu *et al.*, *J. Am. Chem. Soc.* **141**, 17472–17476 (2019).
31. R. D. Astumian, I. Derényi, *Eur. Biophys. J.* **27**, 474–489 (1998).
32. Considering the MW (M_n = 2000) of the PEG, which has around 45 –OCH_2CH_2– repeating units, we estimate that no more than 17 rings can be pumped, ignoring electrostatic interactions, onto the collecting polymer chains (*28*, *29*), as each ring occupies approximately a length on the chain that equates to ~8 atoms.
33. H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew. Chem. Int. Ed.* **40**, 2004–2021 (2001).
34. Y. Wang, M. Frasconi, J. F. Stoddart, *ACS Cent. Sci.* **3**, 927–935 (2017).
35. As demonstrated previously (*28*, *29*), a reduction time of 10 min with Zn dust is more than sufficient to complete the reduction/threading and power the molecular pump. In actual fact, we set the time for the reduction in the present investigation at 30 min in order to achieve exhaustive complexation of a pair of rings onto the pumps. Additionally, taking into consideration the fact that cobaltocene is a stronger reducing reagent than Zn, allowing 30 min for the reduction is more than enough to complete the reduction of all the $BIPY^{2+}$ units.
36. It has been noted that a trace amount of polyrotaxanes with an odd number of rings is present in the ESI-MS spectra

of **S-PR6^{30+}** (fig. S48), **O-PR4^{22+}** (fig. S51), **O-PR8^{38+}** (fig. S55), and **E-PR6^{30+}** (fig. S64).
37. C. Pezzato *et al*., *Angew. Chem. Int. Ed.* **57**, 9325–9329 (2018).
38. ESI-MS failed to provide convincing evidence of the exact number of rings on the electrochemically synthesized **E-PR10^{46+}**. However, the ^{1}H NMR spectrum of **E-PR10^{46+}** (fig. S21) reveals a structure similar to that of **O-PR10^{46+}** (fig. S13), which is obviously different from that of its starting material **E-PR8^{38+}** (fig. S20).
39. G. Ragazzon, L. J. Prins, *Nat. Nanotechnol.* **13**, 882–889 (2018).
40. R. D. Astumian, *Nat. Commun.* **10**, 3837 (2019).
41. R. D. Astumian, B. Robertson, *J. Am. Chem. Soc.* **115**, 11063–11068 (1993).
42. S. Erbas-Cakmak *et al*., *Science* **358**, 340–343 (2017).
43. G. Ragazzon *et al*., *Nat. Nanotechnol.* **10**, 70–75 (2015).

ACKNOWLEDGMENTS

We thank Y. Zhang for help with NMR measurements, A. Gaisin for help with SEC measurements, and T. Kosikova, H. Sun, X. Zhou, and W. Cao for discussions. **Funding:** We thank Northwestern University for its continued support of this research. B.S. and X.L. were supported by NIH grant R01GM128037. This work made use of the Integrated Molecular Structure Education and Research Center (IMSERC) at Northwestern University, which has received support from the Soft and Hybrid Nanotechnology Experimental (SHyNE) Resource (NSF ECCS-1542205), the State of Illinois, and the International Institute for Nanotechnology (IIN). The Bruker 600 MHz NMR spectrometer used in this work is partially funded by NIH grant 1S10OD012016-01. This work made use of the Keck-II facility of Northwestern University's NU*ANCE* Center, which has received support from the SHyNE Resource (NSF ECCS-1542205), the MRSEC program (NSF DMR-1720139) at the Materials Research Center, IIN, the Keck Foundation, and the State of Illinois through IIN. Use of resources of the Keck Biophysics Facility was supported in part by National Cancer Institute award CCSG-P30-CA060553 to the Robert H. Lurie Comprehensive Cancer Center at Northwestern University. **Author contributions:** Y.Q., C.P., C.C., and J.F.S. conceived the study; Y.Q. and J.F.S. designed experiments; Y.Q. conducted the synthesis, NMR, CV, DLS, SEC, and zeta potential analysis; B.S. and X.L. conducted ESI-MS measurements and analysis; D.S. contributed to the graphical design in figures; W. Liu conducted ITC measurements, NMR titration experiments, and NMR analysis; C.P. and Y.F. provided part of the synthetic precursors; C.P., D.S., W. Liu, L.Z., Y.F., Q.-H.G., K.C., W. Li, H.C., M.T.N., Y.S., and R.D.A. commented on the data and all authors contributed to data analysis; and Y.Q. and J.F.S. wrote the draft with input from all authors. **Competing interests:** Y.Q. and J.F.S. have a patent application lodged with Northwestern University (INVO Reference No. NU 2020-106) based on this work. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1247/suppl/DC1
Materials and Methods
Figs. S1 to S87
Tables S1 to S29
References (*44*, *45*)

20 February 2020; resubmitted 5 April 2020
Accepted 27 April 2020
10.1126/science.abb3962

PROTEIN FOLDING

Disordered proteins follow diverse transition paths as they fold and bind to a partner

Jae-Yeol Kim and Hoi Sung Chung*

Transition paths of macromolecular conformational changes such as protein folding are predicted to be heterogeneous. However, experimental characterization of the diversity of transition paths is extremely challenging because it requires measuring more than one distance during individual transitions. In this work, we used fast three-color single-molecule Förster resonance energy transfer spectroscopy to obtain the distribution of binding transition paths of a disordered protein. About half of the transitions follow a path involving strong non-native electrostatic interactions, resulting in a transition time of 300 to 800 microseconds. The remaining half follow more diverse paths characterized by weaker electrostatic interactions and more than 10 times shorter transition path times. The chain flexibility and non-native interactions make diverse binding pathways possible, allowing disordered proteins to bind faster than folded proteins.

The transition path is the molecular trajectory taken by reactants and products in a chemical reaction to cross the free-energy barrier separating them. Although the transition path contains all of the structural information about the process, it has been experimentally studied only recently using single-molecule spectroscopy because barrier crossing is stochastic and thus cannot be synchronized for ensemble measurements and because transition path times are very short (*1*–*3*). To date, transition paths have been studied for folding of proteins (*1*, *3*) and nucleic acids (*4*, *5*) and binding of disordered proteins (*6*, *7*) by using either two-color single-molecule fluorescence or single-molecule force spectroscopy, both of which provide information along only a one-dimensional (1D) reaction coordinate (single distance). Theory and atomistic molecular dynamics simulations, however, predict that individual transition paths for macromolecular conformational changes such as protein folding are very different from reaction pathways for small-molecule chemical reactions (*8*–*12*), even though the kinetics have been successfully described on a 1D free-energy surface (*9*, *13*, *14*). Transition paths are predicted to be more diverse, and there is also much greater structural complexity. These properties are difficult to probe using the 1D measures of single-molecule experiments. In this work, we developed three-color Förster resonance energy transfer (FRET) spectroscopy (*15*–*18*) to obtain multidistance structural information for the transition paths of binding and folding of a disordered protein.

We investigated the binding of the transactivation domain (TAD), the disordered N-terminal domain of the tumor suppressor protein p53, and one of its binding partners, the nuclear coactivator binding domain (NCBD) of the CREB binding protein (*19*). Disordered TAD folds when it binds to its partner (coupled folding and binding) (*20*). Previously, we have shown that binding of TAD and NCBD is diffusion limited (*6*); two molecules associate almost every time they encounter each other. Because TAD is disordered before binding and there is no orientational preference when the two molecules approach, this fast association suggests that binding should take place along diverse transition paths. Otherwise, most encounters would result in rapid dissociation before folding of TAD. The goal of the current work is to test this hypothesis experimentally and extract the binding mechanism from the quantitative characterization of the distribution of the transition path, including the conformational diversity, fraction, and transition path times of the major clusters of pathways.

For the three-color FRET experiment, we attached the donor (D; Alexa Fluor 488) and acceptor 1 (A1; Alexa Fluor 594) to the N and C termini of TAD, respectively, and attached acceptor 2 (A2; CF680R) to the NCBD so that we could probe the end-to-end distance change of TAD (i.e., conformational change) and simultaneously monitor interactions between NCBD and the two ends of TAD (Fig. 1A and fig. S1). A simplified illustration of diverse transition paths and ideal acceptor fraction (i.e., fraction of acceptor photons in a bin) trajectories is shown in Fig. 1 as an example. The structural mechanism of the process can be obtained by experimentally measuring the distribution of these transition paths in terms of the order of interactions between TAD and NCBD, similar to the order of helix formation in protein folding simulation studies (*9*–*11*, *21*). The pathways would be distinguishable by different patterns of the ideal acceptor fraction trajectories in three-color FRET (Fig. 1B), whereas the information obtained by two-color FRET is limited. For example, the average distance between D and A1 is similar for all three pathways. In practice, however, the photon count rate is limited and the transition paths are not well resolved in the experimental binned trajectories, even with the bin time of 200 μs, collected at high illumination intensity (photon count rate of ~400 ms^{-1}; Fig. 2A). Thus, transitions look almost instantaneous. To enhance the

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA.
*Corresponding author. Email: chunghoi@niddk.nih.gov

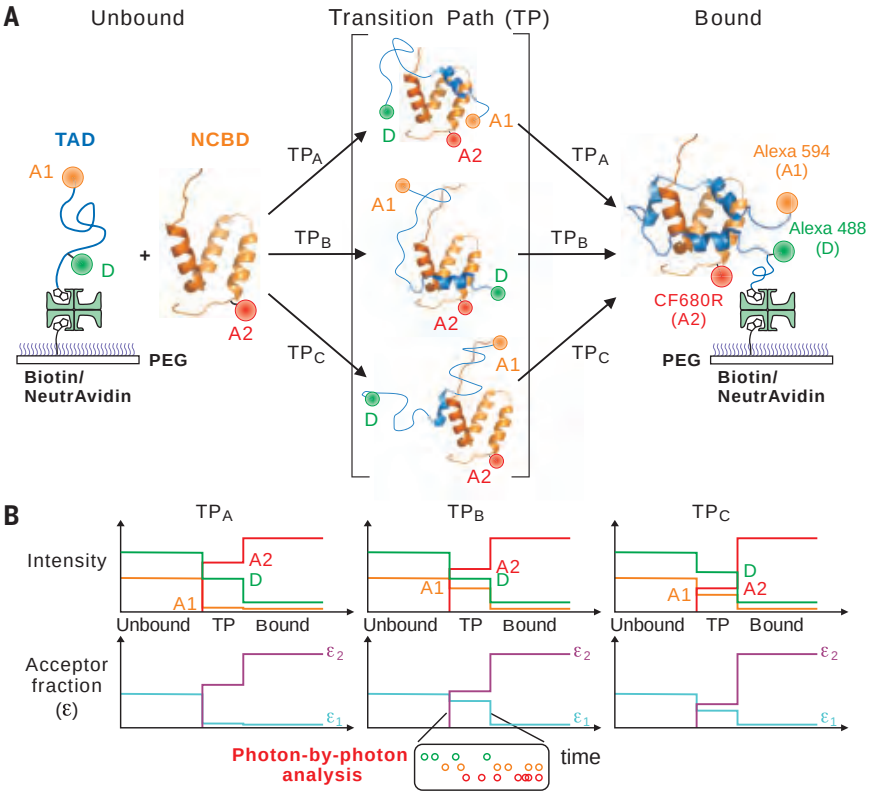


Fig. 1. Three-color FRET to probe diverse binding transition paths. (**A**) Multiple binding transition paths. The donor (D)– and acceptor 1 (A1)–labeled TAD is immobilized on a polyethylene glycol (PEG)–coated glass surface. Unstructured TAD folds upon binding acceptor 2 (A2)–labeled NCBD in solution. Illustrated are three transition paths—TP_A, TP_B, and TP_C—along which the C-terminal, N-terminal, and middle helices of TAD, respectively, fold first, followed by folding of the other part of the molecule. The real transition paths should be much more complex and diverse. In this work, we found two major clusters of transition paths with extra diversity in each cluster. (**B**) Idealized binned (i.e., infinite number of photons in a bin) trajectories of fluorescence intensity and fractions of acceptors demonstrating binding transitions along the three transition paths in (A). Acceptor fractions are defined as $\varepsilon_1 = n_{A1}/n$ and $\varepsilon_2 = n_{A2}/n$, where n_{A1}, n_{A2}, and n are A1, A2, and total photon count rates, respectively. The patterns of fluorescence intensities and acceptor fractions for the three transition paths are different. A maximum likelihood method was used to analyze photon trajectories without binning (photon-by-photon analysis) and to determine the acceptor fractions and duration of the transition paths.

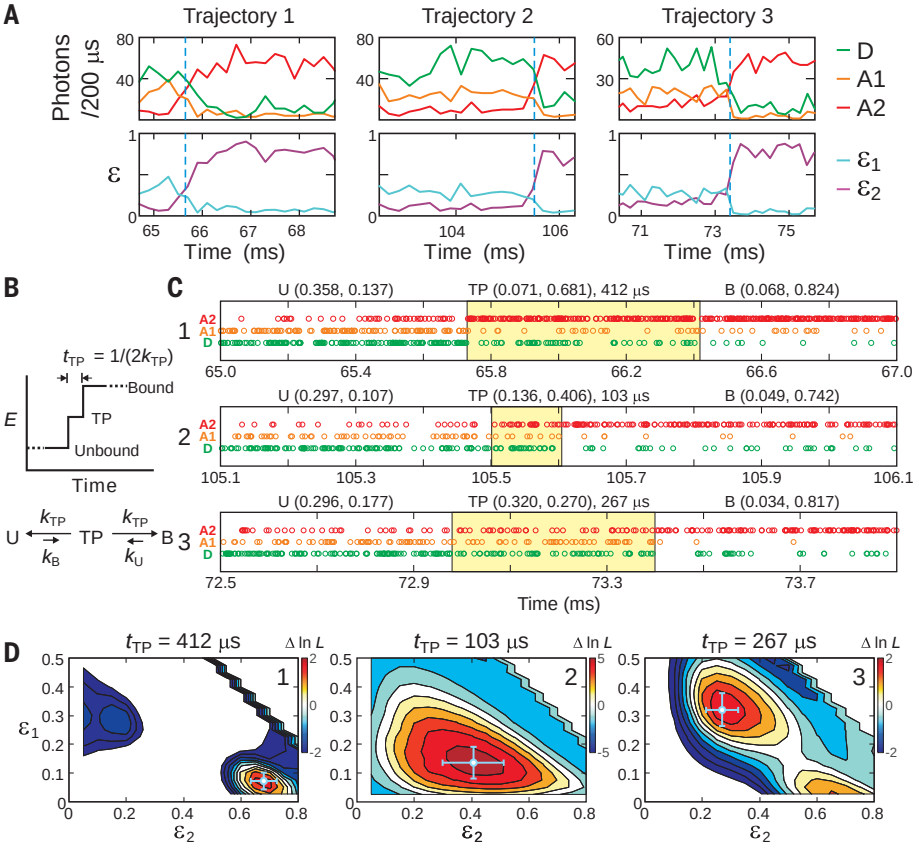


Fig. 2. Maximum likelihood analysis of individual transition paths. (**A**) Representative three-color binned trajectories (200-μs bin time) of fluorescence intensity and acceptor fractions during binding transitions at 0 mM NaCl. (**B**) One-step transition model. The transition path is discretized and analyzed using the three-state model (B, bound; U, unbound; E, FRET efficiency; k, rate constant). The lifetime and acceptor fractions of the intermediate state correspond to the transition path time and average acceptor fractions of the transition paths. (**C**) Three-color photon trajectories near the transitions of the three trajectories in (A). Values in parentheses are the transition path times and ε_1 and ε_2 of the corresponding states obtained from the maximum likelihood method. Yellow shaded regions indicate transition path windows, as determined from the Viterbi algorithm using the maximum likelihood parameters (see supplementary materials). (**D**) 2D likelihood plots of the three transitions in (C). $\Delta\ln L(\varepsilon_1, \varepsilon_2, t_{TP}) = \ln L(\varepsilon_1, \varepsilon_2, t_{TP}) - \ln L(0)$ is the difference of the log likelihood values between the three-state model and the two-state model (i.e., instantaneous transition). $\Delta\ln L$ values of the white area outside the likelihood peaks are smaller than the minimum value of the color bar. Error bars indicate SDs obtained from the diagonal elements of the covariance matrix calculated at the maximum of the likelihood function.

time resolution, we analyzed photon trajectories without binning by using a powerful maximum likelihood method (*22*).

The kinetic model used to describe binding along the transition path is shown in Fig. 2B and fig. S2. Individual transitions were analyzed with the single-path (three-state) model (Fig. 2B), in which the transition path is approximated to be a single intermediate state between the bound and unbound states. The transition path is therefore characterized by the lifetime (t_{TP}) and fractions of the A1 (ε_1) and A2 (ε_2) count rates (see eq. S3 and S5 in the supplementary materials for the definition of acceptor fractions). The likelihood (L) is a function of these variables [$L \equiv L(\varepsilon_1, \varepsilon_2, t_{TP})$], and the parameters most consistent with the data can be obtained by maximizing L (see the supplementary materials for the details of models and likelihood analysis). More complex models such as a multistep transition may be used to capture more realistic gradual changes in the transition path if the photon count rate is sufficiently high.

The individual photon trajectories (Fig. 2C and figs. S3 to S5) directly demonstrate heterogeneous transition paths. The transition path (yellow shaded region) of the top trajectory in Fig. 2C (Trajectory 1) shows that the A1 count rate is the lowest and that the A2 count rate is higher than the D count rate. This pattern indicates that the C-terminal side of TAD, where A1 is attached, interacts with NCBD, similar to TP_A (transition path A) in Fig. 1A. The corresponding likelihood peak appears at (ε_2, ε_1) = (0.68, 0.07) in Fig. 2D (left). On the other hand, count rates of all D, A1, and A2 photons are comparable for the transition path of the bottom trajectory (Trajectory 3), and the likelihood peak appears at (ε_2, ε_1) = (0.27, 0.32) (Fig. 2D, right).

Figure 3A shows the overall individual analysis of three-color trajectories. When the photon count rate is not high enough for the transitions with short t_{TP}, the uncertainty in the determination of the acceptor fractions is large. Therefore, the transitions with relatively high statistical significance—determined by the log-likelihood difference between the two-state and three-state models being greater than 1 [$\Delta \ln L > 1$)]—are plotted in Fig. 3A (transition data with different $\Delta \ln L$ cutoff values are also plotted in fig. S6). The overall transition data show highly diverse transition paths in terms of acceptor fraction values as well as transition path times. The transitions were clustered into two groups by using a clustering algorithm described in the supplementary materials. The transitions with long t_{TP} (>200 μs, orange and red circles) are localized at (ε_2, ε_1) = (0.65, 0.07) (see fig. S7A for individual plots of four different t_{TP} ranges for clarity of localization). We denote this cluster of transition path as TP1 (area inside green dashed lines in Fig. 3A). On the other hand, t_{TP} of the transitions appearing in the middle of the 2D plot, denoted as TP2 (area inside purple dashed lines), is much shorter (green and blue circles) (see fig. S7B for separate plots of the two clusters for clarity of the distributions). Transitions with $\varepsilon_2 <$ 0.25 are not included in the further analyses because these transitions likely result from statistical fluctuations of the photon count rates, as transitions with acceptor fractions in this range are also observed in the simulations (see also figs. S9 to S11). The order-of-magnitude difference in the transition path times indicates that TP1 and TP2 are clearly different kinds of transition paths.

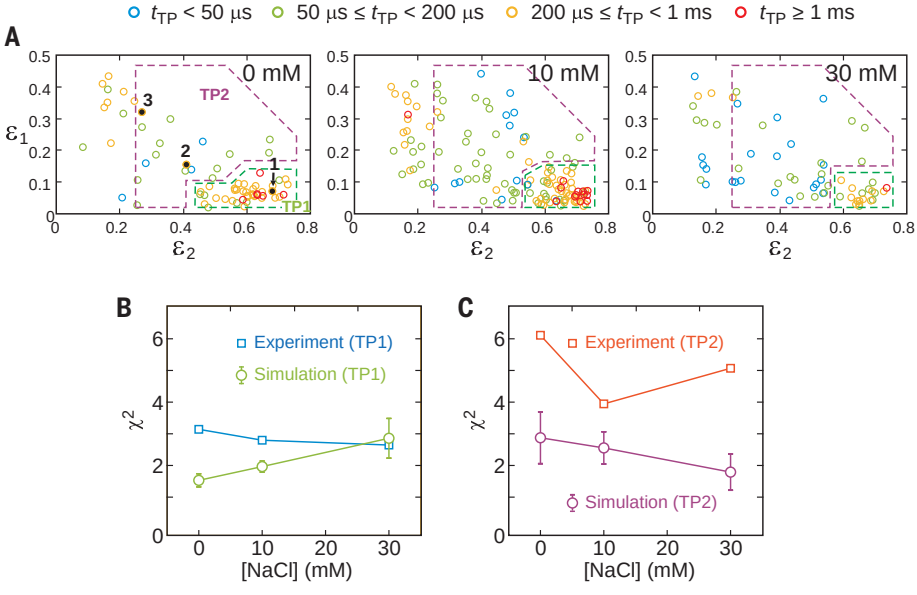


Fig. 3. Analysis of individual transition paths. (**A**) Acceptor fractions of individual transition paths from three-color trajectories at 0, 10, and 30 mM NaCl. Transitions with $\Delta \ln L > 1$ are plotted. Transition path times are indicated by four different colors: blue (t_{TP} < 50 μs), green (50 ≤ t_{TP} < 200 μs), orange (200 ≤ t_{TP} < 1000 μs), and red (t_{TP} ≥ 1000 μs). See fig. S7A for individual plots of four different t_{TP} ranges. The transitions were clustered into two groups, TP1 and TP2, surrounded by green and purple dashed lines, respectively, which are used for the further analyses in (B) and (C). TP1 and TP2 clusters are also plotted separately in fig. S7B for clarity. In the 0 mM plot, black filled circles indicate the three transitions shown in Fig. 2. (**B** and **C**) Comparison of χ^2 values of transitions in the experiment (squares) and the double–transition path simulation (circles) for (B) TP1 and (C) TP2. Error bars are SDs calculated from five simulations. See the supplementary materials for cluster analysis and calculation of χ^2.

In addition to the three-color trajectories, we also analyzed two-color trajectories. The results of the two-color experiments with three different pairs of labeling positions are consistent with those of the three-color experiment (see supplementary materials and figs. S3 to S5).

Although we categorized the transitions into two clusters (TP1 and TP2), this does not mean that there are only two pathways. The TP2 distribution is much more scattered than the TP1 distribution, which suggests that TP2 may consist of more than one cluster of transition paths. However, this apparent diversity could also result from the larger uncertainty of the acceptor fractions due to shorter t_{TP} and the small number of photons collected per transition.

To test whether the broad distribution of the experimentally measured TP1 and TP2 results from the real diversity of the transition path or from photon statistics fluctuations, we compared the diversity parameter χ^2 (eq. S8) of the experimental data and simulation. To perform the simulation, we first fit the entire experimental data with both single-path (i.e., TP1; fig. S2A) and double-path (i.e., TP1 and TP2; fig. S2B) models (fitting parameters are shown in Fig. 4 and tables S1 to S3) and used these maximum likelihood parameters to recolor the experimental photon trajectories (see supplementary materials for the recoloring procedure). The 2D plots of acceptor fractions of individual TPs of three-color trajectories of five simulations for single- and double-path models are shown in figs. S10 and S11. DA2 (donor and A2) simulation results are also displayed in figs. S12 and S13. The χ^2 values (eq. S8) were then calculated for the transitions in TP1 and TP2. If the χ^2 values of TP1 or TP2 from the experiment are significantly larger than the values from the double-path simulation, the experimental distribution of the acceptor fractions in that cluster likely results from more than a single transition path.

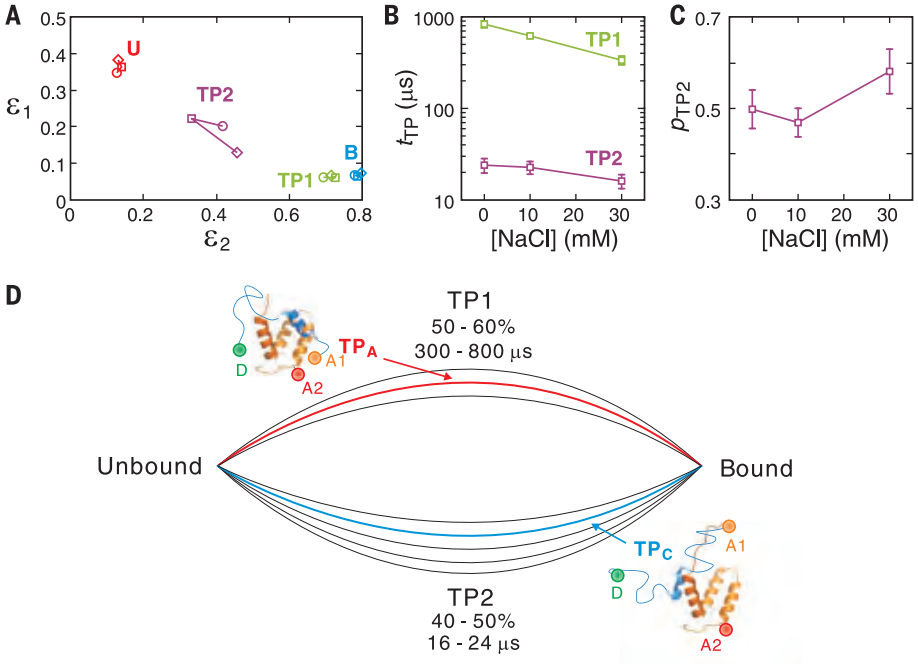


Fig. 4. Collective maximum likelihood analysis of binding transition paths using the double-path model. Parameters were determined from the global analysis of transitions in the three-color and two-color photon trajectories. (**A**) Acceptor fractions for three-color trajectories. Values at 0, 10, and 30 mM NaCl are indicated by circles, squares, and diamonds, respectively. (**B**) Transition path times of TP1 and TP2. (**C**) Fraction of TP2 (p_{TP2}). (**D**) Illustration of diverse transition paths connecting the bound and unbound states, according to experimental measurement of the diversity, t_{TP}, and fractions of TP1 and TP2. TP1 includes TP_A (red), where the C-terminal side of TAD strongly interacts with NCBD and TAD is compact; TP2 includes TP_C (blue), where the middle of TAD interacts with NCBD. More lines in TP2 indicate that it consists of more diverse transition paths than TP1.

Indeed, the experimental χ^2 values of TP1 at 0 and 10 mM NaCl are larger than the simulation values (Fig. 3B). The value at 0 mM is the largest, presumably owing to the broad distribution of the TP1 transitions along the ε_2 axis compared with that of other NaCl concentrations (Fig. 3A); this indicates more diversity in TP1 at lower ionic strength. Compared with the TP1 values, the experimental χ^2 values of TP2 are more than twice as large as the simulated values at 0 and 30 mM, although the difference is small at 10 mM NaCl (Fig. 3C). This result shows that TP2 is much more diverse than TP1.

In addition to the diversity of the binding transition path, another characteristic of intrinsically disordered protein binding is relatively long transition path times (10 μs to 1 ms; Fig. 3) compared with those of protein folding (1 to 10 μs) (*1*, *3*). Because transition paths can be clustered into two representative paths, TP1 and TP2, we analyzed the total likelihood of two- and three-color trajectories globally with the double-path model (fig. S2B) to determine the average transition path times of the two clusters more accurately. The fitting parameters are summarized in Fig. 4 and table S3 (parameters were also used for the recoloring simulation described above). The transition path time of TP1 (300 to 800 μs) is more than an order of magnitude longer than that of TP2 (16 to 24 μs) (Fig. 4B and table S3). t_{TP} of TP1 becomes shorter as NaCl concentration is increased (Fig. 4B), consistent with our previous two-color FRET result (*6*) (equivalent to DA1 trajectory analysis in the current work with a different acceptor fluorophore). This result indicates that the transient complex in TP1 is stabilized by strong electrostatic interactions. The very low A1 fraction (ε_1) of TP1 indicates that the C terminus of TAD, where A1 is attached, interacts strongly with NCBD (TP_A in Fig. 1A) so that most of the A1 excited energy is transferred to A2. The C-terminal electrostatic interactions should result from the large negative net charge (−8) localized in the C-terminal side of TAD (fig. S1). Because the binding interface of the bound complex is mainly hydrophobic (*19*), these electrostatic interactions must be non-native (*6*). The unusually long transition path time of TP1 relative to that of protein folding would be required for escaping from this potentially misfolded state by rearranging and proper folding of TAD through breaking these non-native (albeit strong) interactions.

Three-color FRET makes it possible to observe a new cluster of transition paths (TP2) with short transition path times. From the intermediate and diverse acceptor fraction values, some of the transition paths in TP2 likely resemble TP_B and TP_C. t_{TP} of TP2 transitions (16 to 24 μs) is much closer to the folding transition path time of a designed protein, α_3D (12 μs) (*23*, *24*), suggesting fewer electrostatic interactions compared with TP1. However, the folding transition path time of α_3D is almost an order of magnitude longer than that of naturally evolved proteins (*25*), owing to the formation of non-native salt bridges during folding (*24*). This result suggests that some non-native electrostatic interactions remain in TP2. The decrease of t_{TP2} with increasing NaCl concentration (Fig. 4B) supports this interpretation. The relatively strong and weak electrostatic interactions in TP1 and TP2 would result in the increased fraction of TP2 at a higher NaCl concentration, although the increasing trend in Fig. 4C is not so statistically significant over the small range of NaCl concentration. Signatures of the shift of the binding mechanism by mutation, such as the modulation of dominant binding pathways (*26*) and variation of the transition state ensemble (*27*), have also been observed for other disordered protein binding systems.

In the context of our results (summarized in Fig. 4D), one might ask: What is the advantage of diverse transition paths with long t_{TP}? Diverse pathways can increase the binding rate of disordered proteins. Unlike the association of two folded proteins, in which orientational constraints can reduce the association rate substantially (*28*, *29*), non-native interactions between a flexible disordered chain and a binding partner would allow binding to be initiated from diverse conformations and relative orientations (*30*). The long transition path time is required for converting non-native interactions to native interactions by rearranging and folding of an intrinsically disordered protein. Because the rearrangement and folding occur with the two molecules being held together by non-native interactions without dissociation, the long transition path time would not reduce the rate of association. Thus, extremely fast association, reaching the regime of diffusion-limited reactions, is possible for the TAD and NCBD (*6*). It should be noted that, in protein folding, non-native interactions slow the folding rate (*24*), whereas in our case they speed up the rate of association.

The diversity of protein folding and unfolding pathways is a central feature of the energy

landscape theory and has been studied for several naturally evolved proteins (*31*–*34*) and indirectly observed for DNA hairpins through variance in transition path shape (*5*). The diverse pathways on an energy landscape funneled toward the native structure enable fast and efficient protein folding—because it would take a long time to find a single, well-defined folding pathway (*8*, *12*, *35*, *36*)—but have not been experimentally measured. Our observation of diverse transition paths in coupled folding and binding of a disordered protein is a direct demonstration of this principle.

REFERENCES AND NOTES

1. H. S. Chung, W. A. Eaton, *Curr. Opin. Struct. Biol.* **48**, 30–39 (2018).
2. N. Q. Hoffer, M. T. Woodside, *Curr. Opin. Chem. Biol.* **53**, 68–74 (2019).
3. H. S. Chung, *J. Mol. Biol.* **430**, 409–423 (2018).
4. K. Neupane *et al.*, *Science* **352**, 239–242 (2016).
5. N. Q. Hoffer, K. Neupane, A. G. T. Pyo, M. T. Woodside, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8125–8130 (2019).
6. J.-Y. Kim, F. Meng, J. Yoo, H. S. Chung, *Nat. Commun.* **9**, 4707 (2018).
7. F. Sturzenegger *et al.*, *Nat. Commun.* **9**, 4708 (2018).
8. J. D. Bryngelson, J. N. Onuchic, N. D. Socci, P. G. Wolynes, *Proteins* **21**, 167–195 (1995).
9. K. Lindorff-Larsen, S. Piana, R. O. Dror, D. E. Shaw, *Science* **334**, 517–520 (2011).
10. R. B. Best, G. Hummer, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3263–3268 (2016).
11. E. R. Henry, R. B. Best, W. A. Eaton, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17880–17885 (2013).
12. W. A. Eaton, P. G. Wolynes, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E9759–E9760 (2017).
13. B. C. Gin, J. P. Garrahan, P. L. Geissler, *J. Mol. Biol.* **392**, 1303–1314 (2009).
14. R. B. Best, G. Hummer, W. A. Eaton, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17874–17879 (2013).
15. S. Hohng, C. Joo, T. Ha, *Biophys. J.* **87**, 1328–1337 (2004).
16. J.-P. Clamme, A. A. Deniz, *ChemPhysChem* **6**, 74–77 (2005).
17. J. Yoo, J. M. Louis, I. V. Gopich, H. S. Chung, *J. Phys. Chem. B* **122**, 11702–11720 (2018).
18. A. Barth, L. Voith von Voithenberg, D. C. Lamb, *J. Phys. Chem. B* **123**, 6901–6916 (2019).
19. C. W. Lee, M. A. Martinez-Yamout, H. J. Dyson, P. E. Wright, *Biochemistry* **49**, 9964–9971 (2010).
20. P. E. Wright, H. J. Dyson, *Curr. Opin. Struct. Biol.* **19**, 31–38 (2009).
21. S. Piana, K. Lindorff-Larsen, D. E. Shaw, *Biophys. J.* **100**, L47–L49 (2011).
22. I. V. Gopich, A. Szabo, *J. Phys. Chem. B* **113**, 10965–10973 (2009).
23. H. S. Chung, W. A. Eaton, *Nature* **502**, 685–688 (2013).
24. H. S. Chung, S. Piana-Agostinetti, D. E. Shaw, W. A. Eaton, *Science* **349**, 1504–1510 (2015).
25. H. S. Chung, K. McHale, J. M. Louis, W. A. Eaton, *Science* **335**, 981–984 (2012).
26. D. Wu, H.-X. Zhou, *Sci. Rep.* **9**, 6172 (2019).
27. A. Toto *et al.*, *Sci. Rep.* **6**, 21994 (2016).
28. M. Schlosshauer, D. Baker, *Protein Sci.* **13**, 1660–1669 (2004).
29. S. Qin, X. Pang, H.-X. Zhou, *Structure* **19**, 1744–1751 (2011).
30. H.-X. Zhou, *Trends Biochem. Sci.* **37**, 43–48 (2012).
31. E. J. Guinn, B. Jagannathan, S. Marqusee, *Nat. Commun.* **6**, 6861 (2015).
32. J. Schönfelder, R. Perez-Jimenez, V. Muñoz, *Nat. Commun.* **7**, 11777 (2016).
33. J. Yoo, J. M. Louis, H. S. Chung, *Biophys. J.* **117**, 1456–1466 (2019).
34. C. Charlier *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4169–E4178 (2018).
35. K. A. Dill, H. S. Chan, *Nat. Struct. Mol. Biol.* **4**, 10–19 (1997).
36. M. Oliveberg, P. G. Wolynes, *Q. Rev. Biophys.* **38**, 245–288 (2005).

ACKNOWLEDGMENTS

We thank W. A. Eaton, A. Szabo, I. V. Gopich, R. B. Best, J. Yoo, and F. Meng for numerous helpful discussions and comments; I. V. Gopich for advice on the cluster analysis; P. G. Schultz for sharing the plasmid for the expression and incorporation of the unnatural amino acid; and J. M. Louis for advice and suggestions on protein expression and purification. **Funding:** This work was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases, NIH. **Author contributions:** J.-Y.K. and H.S.C. designed the research and wrote the manuscript. J.-Y.K. performed the research and data analysis. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. Materials used in this study are available from H.S.C. under a material transfer agreement with the National Institutes of Health.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1253/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S13
Tables S1 to S4
References (*37*–*47*)

27 November 2019; accepted 10 April 2020
10.1126/science.aba3854

PLASTIC POLLUTION

Plastic rain in protected areas of the United States

Janice Brahney[1]*, Margaret Hallerud[1], Eric Heim[1], Maura Hahnenberger[2], Suja Sukumaran[3]

Eleven billion metric tons of plastic are projected to accumulate in the environment by 2025. Because plastics are persistent, they fragment into pieces that are susceptible to wind entrainment. Using high-resolution spatial and temporal data, we tested whether plastics deposited in wet versus dry conditions have distinct atmospheric life histories. Further, we report on the rates and sources of deposition to remote U.S. conservation areas. We show that urban centers and resuspension from soils or water are principal sources for wet-deposited plastics. By contrast, plastics deposited under dry conditions were smaller in size, and the rates of deposition were related to indices that suggest longer-range or global transport. Deposition rates averaged 132 plastics per square meter per day, which amounts to >1000 metric tons of plastic deposition to western U.S. protected lands annually.

The world produced 348 million metric tons of plastic in 2017, and this number grows every year by ~5% (*1*, *2*). A large proportion of this production accumulates as waste in the environment, and progressive fragmentation leads to the presence of secondary plastics in terrestrial, freshwater, atmospheric, and marine environments (*2*). Extremely high resilience and longevity give plastics their utility, but these same characteristics lead to the unrestrained accumulation of synthetic materials in nearly every ecosystem on the planet (*3*). Though atmospheric microfibers have recently been documented in Europe and the Arctic (*4*, *5*), the route of primary or secondary microplastics (microfibers and particles) to the atmosphere has not been clear. Primary microplastics are defined as plastics that were manufactured in the size range observed (e.g., microbeads), whereas secondary plastics are derived from the fragmentation of larger pieces of plastics through physical abrasion and/or weakening after exposure to ultraviolet light. To determine potential sources of atmospheric microplastics and the rate of their accumulation in conservation areas of the United States, we quantified the fallout of primary and secondary microplastics to 11 remote and protected areas in both wet atmospheric deposition, collected at week-long intervals while precipitation occurred (n = 236 samples), and dry atmospheric deposition, collected at monthly or bimonthly intervals (n = 103). We used relationships between plastic deposition rates and the intersections of air-mass back trajectories with population centers, contemporaneous dust (soil) deposition, global indices of climate, and plastic composition to identify both emission and product commodity sources. Understanding the sources of microplastics to the atmosphere—both in terms of emission points and product commodities—will, in turn, allow us to implement scale-relevant solutions to mitigate plastic pollution.

Microplastics were present in 98% of all of the wet and dry samples analyzed from U.S. protected areas. Observed microplastic particle sizes were between 4 and 188 μm and fiber sizes were between 20 μm and ~3 mm, with average widths and depths of 18 and 6 μm, respectively (fig. S2). Approximately 70% of the particles were within the size range for long-range and even global transport of dust (<25 μm) (*6*, *7*), whereas most fiber lengths suggested regional transport (10 to 1000 km) (*8*). Because plastic density (0.65 to 1.8 g cm^{-3}) is lower than that of soil particles (~2.65 g cm^{-3}) (*9*), microplastics are more transportable. Fibers, in particular, have greater surface area–to–volume ratios, which increase drag forces

[1]Department of Watershed Sciences, Utah State University, Logan, UT 84322, USA. [2]Geosciences Department, Salt Lake Community College, Salt Lake City, UT 84123, USA. [3]Materials and Structural Analysis Division, Thermo Fisher Scientific, San Jose, CA 95134, USA.
*Corresponding author. Email: janice.brahney@usu.edu

and reduce settling velocity. This process may be similar to ballooning in spiders, where a combination of electrostatic forces and drag allows spiders attached to silk fibers to travel thousands of kilometers (*10*).

Daily 48-hour atmospheric back-trajectory analyses were determined using the Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model (*11*, *12*) and were compared with weekly wet plastic deposition rates through 2018. Our analyses suggest that wet-deposited microplastics originate from different source regions than those that are dry deposited. Wet plastic deposition rates at half of the sites were significantly correlated to population metrics, as determined by the intersection of the air mass with population centers (Table 1). Distance traveled, mean wind speeds, and contemporaneous dust deposition also described significant portions of the variance noted at individual sites. We observed that microplastics deposited in wet conditions are larger in size and lower in number (fig. S2) and are correlated to both dust deposition and

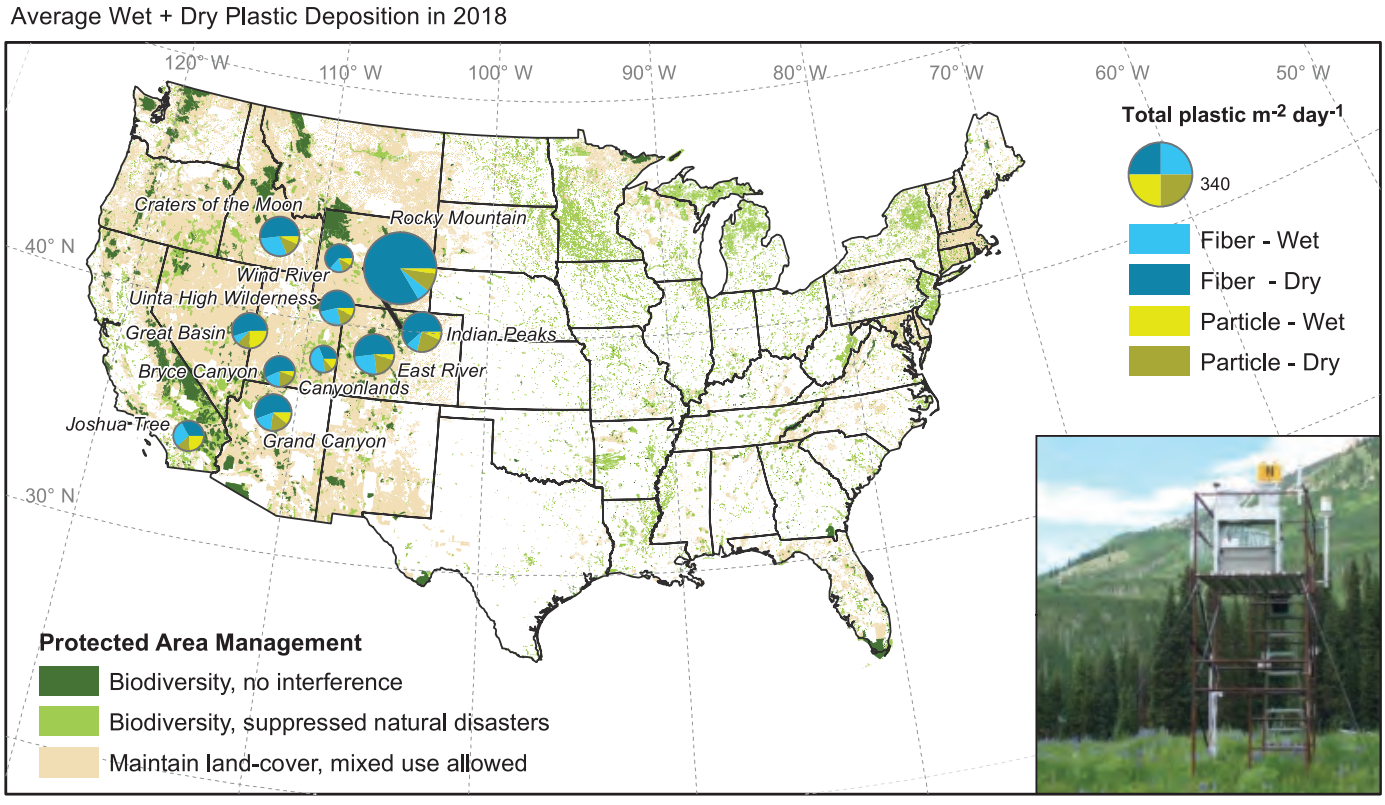


Fig. 1. Average deposition rates of plastic fibers and particles, wet plus dry, to selected national parks and wilderness areas of the United States. The pie chart sizes reflect plastic fluxes to each site. Protected areas base map is from the United States Geological Survey (USGS) Protected Areas Database (PAD). Pictured on the bottom right is a standard National Atmospheric Deposition Program (NADP) Aerochem Metrics wet-dry collector located in East River, Colorado.

Table 1. Relationships between weekly wet plastic deposition rates, dust, population statistics, and air-mass trajectories. Pearson correlation coefficients (r) and model coefficients of determination (r^2) between wet plastic deposition rates and potential drivers. Full model selection is based on the Akaike Information Criterion (AIC), and parameters included are shown in bold. NA, not applicable.

National park or wilderness	Dust (r)	Total population (r)	Total populated area (r)	Distance (r)	Mean wind speed (r)	Full model (r^2)
Grand Canyon, AZ	0.16	0.70***	**0.56****	**0.80*****	0.41	0.69***
Wind River Range, WY	**0.74*****	**0.32**	**0.31**	0.34*	**0.3**	0.77***
Craters of the Moon, ID	−0.11	0.05	**0.43****	**−0.12**	0.01	0.21**
Rocky Mountain, CO	**0.27***	**0.35****	0.12	0.18	0.05	0.20**
Joshua Tree, CA	0.16	**−0.45**	**−0.24**	**0.96****	0.63*	0.71*
Uinta High Wilderness, UT	**0.2**	**−0.48**	**−0.32**	**0.08**	**0.11**	0.86
Canyonlands, UT	**0.44***	0.01	−0.16	0.05	**0.25**	0.19
Indian Peaks, CO	**0.77****	**0.42**	**0.66***	**0.29**	**−0.15**	0.99**
East River, CO	**0.58*****	−0.11	−0.12	−0.05	−0.04	0.34***
Great Basin, NV	**0.41***	**0.51****	**0.17**	**0.48****	0.18	0.59***
Bryce Canyon, UT	−0.13	−0.02	0.001	−0.06	−0.1	NA

*P < 0.1; **P < 0.05; ***P < 0.01.

CREDIT: ERIC HEBERT/ENVIRONMENTAL ENGINEERING & MEASUREMENT SERVICES, INC.

population metrics. This observation reflects the role of regional storms in the entrainment and subsequent rainout of microplastics, as these storms often pass through urban centers or over erodible soils. In contrast, dry deposition shows a negative relationship with regional dust deposition rates and is related instead to indices that represent broad-scale atmospheric patterns, specifically a more southerly jet stream. This suggests that dry-deposited plastics are subject to large-scale, global dispersion (Table 2).

Microfibers made up most of the synthetic material found in both wet (66%) and dry (70%) deposition. Fiber compositions were mainly consistent with those of textiles used for clothing, including cotton, polyester, and nylon. We also observed fibers composed of polyolefin, which is more commonly used for household and vehicle carpeting, as well as polytetrafluoroethylene and polyethylene fibers, which are used in a variety of industrial applications (*13*). Industrial coatings on fibers, such as Valbond 6053, were also identified, which underscores the diversity in microfiber sources to U.S. protected areas. It is worth noting that polypropylene and polytetrafluoroethylene are also commonly used in outdoor gear, including fleece, tents, waterproof clothing, and climbing ropes (*13*). Because microfibers are known to shed from clothing during normal wear (*14*), emissions from park users may contribute to the observed deposition rates, particularly in national parks with high visitation rates. Clothing fibers are also directly released to the atmosphere during laundry drying at rates that are several times the rates at which fibers are released to wastewater during the washing phase (*15*, *16*), and these fibers are then transported to protected areas during times of favorable wind speeds and trajectories (fig. S5).

The polymer compositions of individual plastic particles smaller than 20 μm were more difficult to identify using Fourier transform infrared (FTIR) spectroscopy because of the diffraction limitation of midinfrared light. However, in subsamples, almost all brightly colored particles that fell within our counting criteria were identified as synthetic using FTIR spectroscopy particle mapping in reflection mode, which allows the mass identification of particles in the subsamples. Using this reflectance mapping technique on 32 subsamples, we found that 2.5 to 5% (on average, 4%) of the identifiable dust particles were synthetic polymers. This included particles and fibers that did not meet our visual counting criteria (because they were clear or white), which suggests that our estimates of plastic deposition rates based on counts are conservative (Table 3). Most plastic particle compositions found in our samples can be linked back to industrial applications and coatings. Polyethylene, polypropylene, polyvinyl acetate, and ethylene-acrylic copolymer were also identified. Approximately 30% of the particles were primary plastic microbeads ranging in size from 5 to 30 μm in a wide variety of colors (fig. S1). Primary plastics derived from personal care products have received much attention but are generally larger in diameter (74 to 800 μm) (*16*) than those we observed. Manufacturers of brightly colored microbeads cite primary uses in research and medical applications as well as industrial paints. We identified several pink microbeads as poly(methyl methacrylate) (PMMA), which is used in a variety of industrial paint and coating applications.

Table 2. Comparison of dry and wet plastic deposition rates and their potential drivers. Pearson correlation coefficients (*r*) between weekly wet and monthly dry deposition rates of plastic fibers and particles and indices of regional and broad-scale climate patterns. The temperature anomaly for the Western United States is used here as an index of jet stream location [data: National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information; www.ncdc.noaa.gov/cag/regional/time-series]. ENSO, El Niño–Southern Oscillation; *F*-stat, *F* statistic.

Deposition type	Dust (*r*)	Season (*F*-stat)	ENSO (*r*)	Temperature anomaly (*r*)
Dry deposition (total)	−0.24**	0.63	0.21**	−0.25***
Dry fiber deposition	−0.22**	0.36	0.19*	−0.23**
Dry particle deposition	−0.24**	5.64***	0.29***	−0.36***
Wet deposition (total)	0.37***	3.61**	−0.13*	−0.02
Wet fiber deposition	0.21***	2.91**	−0.12*	−0.03
Wet particle deposition	0.36***	1.71	−0.08	−0.04

*$P < 0.1$; **$P < 0.05$; ***$P < 0.01$.

Table 3. Annual plastic deposition rates to 11 U.S. protected areas. Estimated annual deposition rates of microfibers and plastic particles to national parks and wilderness areas of the United States. Data are based on observed deposition rates to each site from late 2017 to early 2019.

National park or wilderness	State	Size (km^2)	Mean plastic deposition rate (plastics m^{-2} day^{-1})	Metric tons of plastic per year (visual counts)	Metric tons of plastic per year (FTIR proportions)
Grand Canyon	AZ	4926	112 ± 6	10.7–11.9	11.0–21.3
Wind River Range	WY	7252	68 ± 6	9.3–11.1	10.9–22.3
Craters of the Moon	ID	2893	139 ± 10	7.7–8.8	11.5–19.3
Rocky Mountain	CO	1047	435 ± 8	9.4–9.8	4.2–9.0
Joshua Tree	CA	3200	54 ± 2	3.4–3.7	3.7–9.8
Uinta High Wilderness	UT	1849	120 ± 6	4.3–4.8	1.6–2.8
Canyonlands	UT	1366	48 ± 7	1.2–1.5	3.0–6.1
Indian Peaks	CO	311	148 ± 5	0.9–1.0	0.4–1.3
East River	CO	300	140 ± 9	0.8–0.9	0.4–0.9
Great Basin	NV	312	107 ± 5	0.65–0.72	0.4–1.3
Bryce Canyon	UT	145	80 ± 6	0.22–0.26	0.4–0.8
All western protected areas	USA	496,350	132 ± 6	1012–2419	1185–3773

Entrainment to the atmosphere could easily occur for the many industrial coatings and paints that are applied using aerosol sprays, but these may not be the only atmospheric emission sources. Because the density of most microbeads is lower than that of seawater, entrainment could also occur from the surfaces of aquatic systems through aerosolization under turbulent conditions. An analogous process has been shown to aid in the dispersal of algae and other particles across thousands of kilometers (*17*). The dominant size classes of microbeads observed were <20 μm, therefore also subject to global atmospheric dispersal, which indicates that the source of these beads is not necessarily in the continental United States.

First-order estimates of mass deposition rates to each national park and wilderness area were determined using two independent methods. The first method uses the mean deposition rate based on visual count estimates (Fig. 1 and Table 3) and the range of densities observed for the plastics identified (0.92 to 2.2 g cm^{-3}) to calculate the total annual loading of plastic to each protected area. The second method uses FTIR-based estimates of the polymer proportions within our samples. Method 2 estimates are larger but similar to those of method 1 (r = 0.89, where r is the correlation coefficient). Estimated, site-specific annual deposition rates ranged from 48 ± 7 to 435 ± 9 plastics m^{-2} day^{-1}, or 0.22 to 22 metric tons of plastic per year scaled to each park or wilderness area (Table 3). On the basis of these data, we approximate that >1000 tons of plastic from the atmosphere are delivered to western protected areas in the United States, including national parks and wilderness areas, each year. This is equivalent to ~120 to 300 million plastic water bottles.

The finding that microplastics are ubiquitous in the atmosphere and are transported to distant locations has widespread ecological implications. Though the literature is still sparse on the effects of microplastics on terrestrial organisms (*18*), accidental ingestion of plastics by aquatic organisms has been shown to lead to blockages in the intestinal tract causing internal injury, reduced energy, and behavior modifications (*16*, *18*, *19*). In some cases, ingested plastics have been shown to transfer up the food chain (*16*, *18*, *19*). Less is known about the influence of microplastics on microbes, but recent work has suggested that plastics can influence microbial community composition (*20*). This observation leads to key questions about whether plastic-altered microbial communities in receiving terrestrial ecosystems could lead to changes in biogeochemical processes. As plastics accumulate in pristine wilderness, we may anticipate shifts in community composition, possibly leading to declines in biodiversity on the basis of the different tolerances to the physical and toxicological consequences of consuming microplastics. Further, because plastics can influence thermal and hydrologic properties of soils (*21*), changes in the biogeochemical cycling of nutrients in protected environments may also occur with unforeseen consequences. Many of our study locations are mountain environments that tend to have simple food webs and shallow soils (*22*, *23*), which makes them particularly sensitive to perturbations and might lead to an amplified response to microplastic deposition.

To date, only a handful of studies have quantified atmospheric microplastic loading rates to urban and remote settings (*4*, *5*, *24*), and there is a clear, growing need for these types of studies. We show that the intersection of 48-hour air-mass trajectories with, and their proximity to, population centers are coincident with enhanced rates of plastic deposition (up to 14-fold), though a large proportion of the variation is not explained by these local-to-regional factors alone. This result, combined with the size distribution of identified plastics and the relationship to global-scale climate patterns, suggests that plastic emission sources have extended well beyond our population centers and, because of the longevity of plastics, have spiraled through the Earth system. The long-range transport of microplastics, reminiscent of the global dust cycle but distinctly human in origin, is indicative of the ubiquity of the human fingerprint on atmospheric composition; microplastics have the potential to be found far from initial production and source areas.

In highlighting independent life histories for dry versus wet plastic deposition, we provide additional details on the source, transport, and fate of plastics on Earth's surface. Though regional storms were important in delivering larger plastics to national parks, dry deposition accounted for >75% of the plastic mass deposited. This result, along with the relationship of dry deposition to large-scale climate patterns, suggests that although urban centers may be the initial source, plastics accumulate in the atmosphere over longer time periods, are transported long distances, and are deposited during favorable conditions, such as slower air-mass velocities or intersections with mountain ranges. In fact, dry plastic deposition rates showed a significant and positive relationship to elevation (r = 0.69, P < 0.05). However, key questions remain on emission mechanisms and the transport physics of low-density polymers, including atmospheric lifetimes and the role of latitudinal atmospheric circulation patterns. Greater spatial resolution, particularly across latitudinal gradients, and perhaps in situ aircraft-based sampling would provide the data needed to model the atmospheric limb of the global plastic cycle. Identifying the key mechanisms underpinning plastic emissions to the atmosphere is the first step in developing scalable solutions. The consequences to ecosystems are not yet well understood but are inescapable in the immediate future. If the potential dangers posed by environmental microplastics are to be mitigated, both the scale of the solution and the level of cooperation that will be required call on the engagement of the global community.

REFERENCES AND NOTES

1. PlasticsEurope's Market Research and Statistics Group (PEMRG), "Global plastic production from 1950 to 2017 (in million metric tons)" (Statista, 2018); www.statista.com/statistics/282732/global-production-of-plastics-since-1950/.
2. R. Geyer, J. R. Jambeck, K. L. Law, *Sci. Adv.* **3**, e1700782 (2017).
3. C. M. Rochman, *Science* **360**, 28–29 (2018).
4. S. Allen *et al*., *Nat. Geosci.* **12**, 339–344 (2019).
5. M. Bergmann *et al*., *Sci. Adv.* **5**, eaax1157 (2019).
6. N. Mahowald *et al*., *Aeolian Res.* **15**, 53–71 (2014).
7. P. R. Betzer *et al*., *Nature* **336**, 568–571 (1988).
8. C. R. Lawrence, J. Neff, *Chem. Geol.* **267**, 46–63 (2009).
9. N. C. Brady, R. R. Weil, *The Nature and Properties of Soils* (Prentice Hall, 2008).
10. E. L. Morley, D. Robert, *Curr. Biol.* **28**, 2324–2330.e2 (2018).
11. A. F. Stein *et al*., *Bull. Am. Meteorol. Soc.* **96**, 2059–2077 (2015).
12. G. Rolph, A. Stein, B. Stunder, *Environ. Model. Softw.* **95**, 210–228 (2017).
13. J. Preston, "Man-made fibre" (Encyclopædia Britannica, Inc., 2016); www.britannica.com/technology/man-made-fiber.
14. S. A. Carr, *Integr. Environ. Assess. Manag.* **13**, 466–469 (2017).
15. U. Pirc, M. Vidmar, A. Mozer, A. Kržan, *Environ. Sci. Pollut. Res. Int.* **23**, 22206–22211 (2016).
16. K. Duis, A. Coors, *Environ. Sci. Eur.* **28**, 2 (2016).
17. S. V. M. Tesson, C. A. Skjøth, T. Šantl-Temkiv, J. Löndahl, *Appl. Environ. Microbiol.* **82**, 1978–1991 (2016).
18. M. C. Rillig, *Environ. Sci. Technol.* **46**, 6453–6454 (2012).
19. M. Cole *et al*., *Environ. Sci. Technol.* **47**, 6646–6655 (2013).
20. S. Oberbeckmann, M. G. J. Loeder, G. Gerdts, A. M. Osborn, *FEMS Microbiol. Ecol.* **90**, 478–492 (2014).
21. H. S. Carson, S. L. Colbert, M. J. Kaylor, K. J. McDermid, *Mar. Pollut. Bull.* **62**, 1708–1713 (2011).
22. K. A. Moser *et al*., *Glob. Planet. Change* **178**, 77–95 (2019).
23. M. Beniston, *Mountain Environments in Changing Climates* (Routledge, 2002).
24. R. Dris *et al*., *Environ. Pollut.* **221**, 453–458 (2017).

ACKNOWLEDGMENTS

This study was possible because of collaboration with the National Atmospheric Deposition Program (NADP). We thank the NADP executive committee and director as well as P. Strong, G. Weatherbee, G. Sexstone, A. Baldwin, T. Armstrong, S. Belmont, P. Belmont, J. Draper, P. Wilcock, and J. Ranville for their expertise and T. Atwood for comments on the manuscript. **Funding:** This research was supported by the Utah Agricultural Experiment Station and approved as journal paper no. 9348. The research was supported by the Agricultural Experimental Station research grants UTA01421 and UTA01384, NSF grant 1926559, and a U.S. Department of Agriculture Forest Service Agreement to J.B. **Author contributions:** J.B. designed the project, acquired the data, conducted the analyses, and wrote the original draft. M.Hal. and E.H. developed the geographic information system (GIS) models and generated the spatial data and figures. M.Hah. contributed climate data and interpretations. S.S. conducted the FTIR analyses. All authors edited the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1257/suppl/DC1
Materials and Methods
Figs. S1 to S5
Table S1
References (*25*–*29*)

Data S1 to S4

20 September 2019; accepted 22 April 2020
10.1126/science.aaz5819

CELL CYCLE

Temporal integration of mitogen history in mother cells controls proliferation of daughter cells

Mingwei Min[1]*, Yao Rong[1,2], Chengzhe Tian[1], Sabrina L. Spencer[1]*

Multicellular organisms use mitogens to regulate cell proliferation, but how fluctuating mitogenic signals are converted into proliferation-quiescence decisions is poorly understood. In this work, we combined live-cell imaging with temporally controlled perturbations to determine the time scale and mechanisms underlying this system in human cells. Contrary to the textbook model that cells sense mitogen availability only in the G_1 cell cycle phase, we find that mitogenic signaling is temporally integrated throughout the entire mother cell cycle and that even a 1-hour lapse in mitogen signaling can influence cell proliferation more than 12 hours later. Protein translation rates serve as the integrator that proportionally converts mitogen history into corresponding levels of cyclin D in the G_2 phase of the mother cell, which controls the proliferation-quiescence decision in daughter cells and thereby couples protein production with cell proliferation.

Cells convert extracellular mitogen availability into cell decisions by means of mitogen signaling pathways. One central branch is the mitogen-activated protein kinase (MAPK) pathway, in which growth factors bind their receptors at the plasma membrane and activate the Raf-Mek-Erk MAPK cascade. Erk activation leads to the activation of several transcription factors that promote transcription of cyclin D, activation of cyclin-dependent kinases 4/6 and 2 (CDK4/6 and CDK2), and cell cycle entry (Fig. 1A) (*1*). Although signal transduction from the plasma membrane to Erk takes only 3 min (*2*), commitment to the cell cycle occurs only once per cell cycle (Fig. 1A). Therefore, it is unclear how short–time scale MAPK signals control long–time scale cellular proliferation. Early studies based on quiescent cells released from serum starvation led to the textbook model that cells sense mitogen levels in the early G_1 phase before crossing the restriction point (R point), a point marked by the buildup of CDK2 activity after which cells become mitogen independent for the remainder of the cell cycle (*3*–*6*). In cycling cells, by contrast, blocking mitogenic signals during the mother cell G_2, but not during the daughter cell G_1, impedes cell cycle entry (*6*–*8*). This observation led to the model of a G_2-specific window in the mother cell cycle where cells sense mitogen availability (Fig. 1B) (*6*–*8*), and this model is consistent with the recent observation that the proliferation-quiescence decision is already made in late anaphase of the mother cell cycle in unperturbed cells (*6*, *9*). However, in experiments where the MAPK pathway is continuously inhibited, two variables are changing simultaneously—the duration and the cell cycle timing of MAPK inhibition. Therefore, an alternative model to the G_2-specific window is that cells sense the duration of mitogenic signaling throughout the entire mother cell cycle (Fig. 1B).

To assess these two models, we used live-cell imaging of a CDK2 activity sensor (*6*) to follow the commitment to proliferation in conditions where the cell cycle timing and the duration of MAPK inhibition are decoupled. If mitogenic signaling is only required in G_2, we should be able to inhibit the MAPK pathway for the duration of G_1 and S (4 to 12 hours before mitosis), re-enable MAPK signaling at the end of the S phase, and see no effect on proliferation commitment. Alternatively, if cells continuously integrate mitogens during the entire mother cell cycle, MAPK inhibition in any phase would reduce the fraction of proliferating daughter cells (fig. S1A).

In control conditions, ~80% of MCF10A mammary epithelial cells immediately increase CDK2 activity after the completion of mitosis and are committed to the cell cycle (80% $CDK2^{inc}$; Fig. 1C, blue), as has been seen previously (*6*, *10*). The remainder enter a transient state of quiescence characterized by low CDK2 activity, where they can remain for the rest of the experiment ($CDK2^{low}$; Fig. 1C, orange), or they emerge from this transient quiescence by building up CDK2 activity to reenter the cell cycle ($CDK2^{emerge}$; Fig. 1C, green). To test the two models, we reversibly inhibited the MAPK pathway using a Mek inhibitor (Meki) in asynchronously cycling MCF10A cells for a fixed duration (1, 3, 6, or 9 hours) before washing out the inhibitor. We then computationally grouped cells that received the inhibition in different cell cycle phases on the basis of the time of drug addition relative to anaphase (fig. S1, B and C), and we examined the fraction of proliferating daughter cells in each group by monitoring CDK2 activity (Fig. 1D). We confirmed that Meki does not notably alter the length of the ongoing cell cycle (fig. S1, D and E) and that adding in and washing off Meki rapidly modulates Erk activity, as has been shown previously (fig. S2A) (*11*, *12*). Consistent with previous observations (*6*, *8*), continuous treatment of Meki ("till end"; Fig. 1E) beginning at least 6 hours before mitosis completely blocks proliferation of daughter cells (0% $CDK2^{inc}$).

Notably, mother cells receiving a pulse of Meki as short as 1 hour in length at any time during their cell cycle generate fewer $CDK2^{inc}$ daughter cells (65% $CDK2^{inc}$; Fig. 1E and fig. S2, B and C). This indicates that Mek activity is required throughout interphase in mother cells for efficient daughter cell proliferation and that cells carry the memory of a 1-hour absence in Mek activity all the way through the mother cell cycle and into the daughter cell cycle. The reduction of the $CDK2^{inc}$ fraction coincides with an increase of the $CDK2^{emerge}$ fraction, which suggests that these daughter cells need extra time to commit to the cell cycle (Fig. 1F and fig. S2, B to E). Although the fraction of $CDK2^{inc}$ daughter cells does not vary with the cell cycle timing of the Meki treatment, it decays with the duration of treatment (Fig. 1, E and F, and fig. S2, C and E). The same treatment duration–dependent effect on proliferation is also seen with Erk inhibition (fig. S2F), and upon withdrawal of epidermal growth factor (fig. S2G). Taken together, these results show that MAPK signaling is temporally integrated throughout the mother cell interphase to guide the proliferation of daughter cells.

We next examined where the integration of MAPK signaling occurs in the pathway. MAPK signaling promotes cell cycle entry by means of the activation of cyclin D–CDK4/6 (*13*–*16*). We therefore tested whether CDK4/6 activity is also temporally integrated throughout the entire mother cell cycle. We used palbociclib to inhibit CDK4/6 (CDK4/6i) for 1, 3, 6, or 9 hours or continuously. Contrary to Meki, CDK4/6i shows a cell cycle phase–dependent effect on proliferation—CDK4/6i covering 0 to 3 hours after anaphase maximally blocks proliferation, whereas treatments that cover other time windows show no effect on proliferation (Fig. 1G and fig. S3). This result reveals that cells only require CDK4/6 activity from anaphase until after CDK2 activation (fig. S3). Thus, the integration of MAPK signaling occurs upstream of CDK4/6 (Fig. 1H).

We next set out to determine the molecular nature of the integrator. We reasoned that the integrator should satisfy three criteria: (i) regulate proliferation; (ii) sense a 1- to 3-hour lapse of MAPK activity throughout the cell cycle; and (iii) carry memory of MAPK inhibition (fig. S4A). As first candidates, we considered CDK inhibitor proteins

[1]Department of Biochemistry and BioFrontiers Institute, University of Colorado Boulder, Boulder, CO 80303, USA. [2]Department of Molecular, Cellular and Developmental Biology, University of Colorado Boulder, Boulder, CO 80303, USA.
*Corresponding author. Email: sabrina.spencer@colorado.edu (S.L.S.); min_mingwei@grmh-gdl.cn (M.M.)

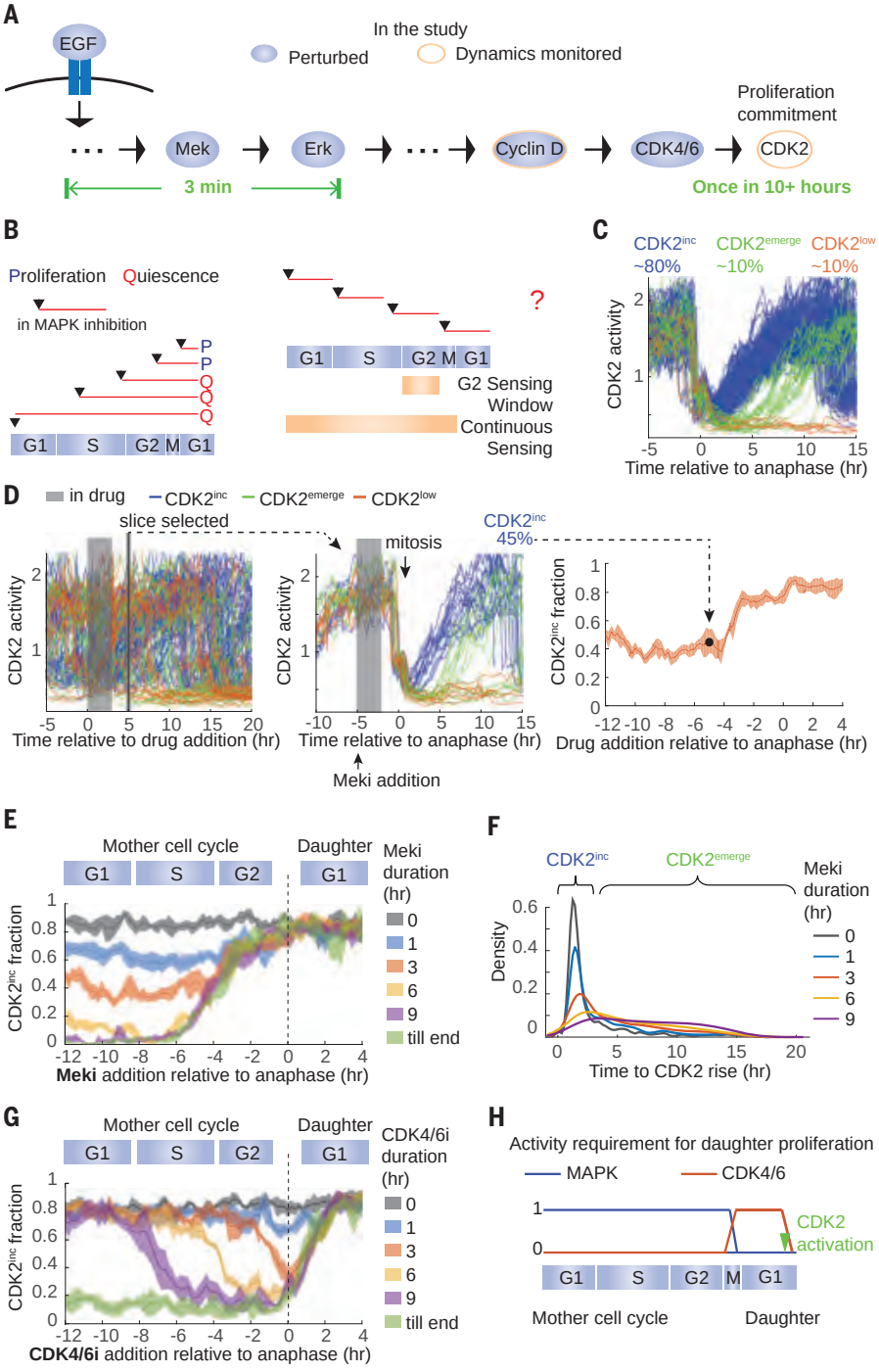


Fig. 1. MAPK activity is temporally integrated throughout the mother cell cycle to control daughter cell proliferation, whereas CDK4/6 activity is only required in early G_1. (**A**) Transduction of mitogen signals to cell cycle machinery. EGF, epidermal growth factor. (**B**) (Left) Depiction of previous MAPK inhibition experiments. (Right) Two potential models of mitogen sensing: Cells sense mitogen availability only in the mother cell G_2 or continuously throughout the mother cell cycle. (**C**) Typical cell CDK2 activity in optimal growth conditions. hr, hour. (**D**) Data processing for establishing daughter cell fate as a function of the time of drug addition, read out as CDK2 activity. (Left) CDK2 activity in asynchronously cycling cells. In this example, cells were treated with Meki for 3 hours before the drug was washed off. (Middle) Traces from cells that underwent mitosis 5 hours after drug addition were extracted and the fraction of $CDK2^{inc}$ daughter cells was calculated. (Right) Repeating this process for every time slice generates the plot of fraction of $CDK2^{inc}$ daughter cells versus the time of drug addition relative to anaphase. (**E**) Fraction of $CDK2^{inc}$ daughter cells treated with 0, 1, 3, 6, or 9 hours Meki or Meki until the end (till end) of the experiment, at various times relative to anaphase. (**F**) Density distribution of the time between anaphase and the rise of CDK2 activity in $CDK2^{inc}$ and $CDK2^{emerge}$ daughter cells. Cells were treated with Meki for the indicated durations starting from the G_1 phase of the mother cell cycle. Areas under the curves were normalized to 1. (**G**) Same as in (E) except with CDK4/6i treatment. (**H**) Summary of the data. All data are from MCF10A cells; all $CDK2^{inc}$ fractions are plotted as means ± 95% confidence intervals shown as shaded bands, where nonoverlapping shading indicates a statistically significant difference as determined by *t* test, with $P < 0.05$.

p21 and p27 (*17*) and cyclin D, the nexus between MAPK signaling and cell cycle entry (*13*, *14*). The inhibitor protein p21 does not satisfy criterion (ii) (fig. S4B), and p27 does not satisfy criteria (i) and (iii) (fig. S4, C to E). Cyclin D satisfies criterion (i)—knockdown of all three cyclin D genes in mother cells impairs proliferation of daughter cells (20 to 30% $CDK2^{inc}$; Fig. 2A and fig. S5, A and B), and overexpression of cyclin D1 rescues the Meki-induced proliferation defect (fig. S5, C and D). We and others have reported that cyclin D1 protein levels increase in the G_2 phase of the cell cycle (*10*, *18*, *19*) (Fig. 2B). This G_2 rise of cyclin D1 protein is attenuated by prior Mek inhibition in a duration-dependent (Fig. 2B, left) and cell cycle phase–independent manner (Fig. 2B, right), satisfying criterion (iii) that cyclin D1 protein levels in G_2 carry memory of MAPK inhibition. However, cyclin D1 protein levels do not sense instantaneous Mek activity because Mek inhibition fails to rapidly alter cyclin D1 protein levels in G_1 and only later blocks cyclin D1 rise in G_2 (Fig. 2C). Thus, criterion (ii) is not satisfied, and cyclin D1 is not the integrator. Nonetheless, these results reveal that the integrator relays the information of MAPK activity in the mother cell G_1 to cyclin D protein levels in G_2 (fig. S4A).

It has long been established that the MAPK pathway regulates cyclin D transcription. However, the five proposed cyclin D transcription factors—Ets1, Fos, Fra1, Jun, and Myc (*20*–*22*)—either do not satisfy the three criteria above or do not rescue the Meki-induced proliferation

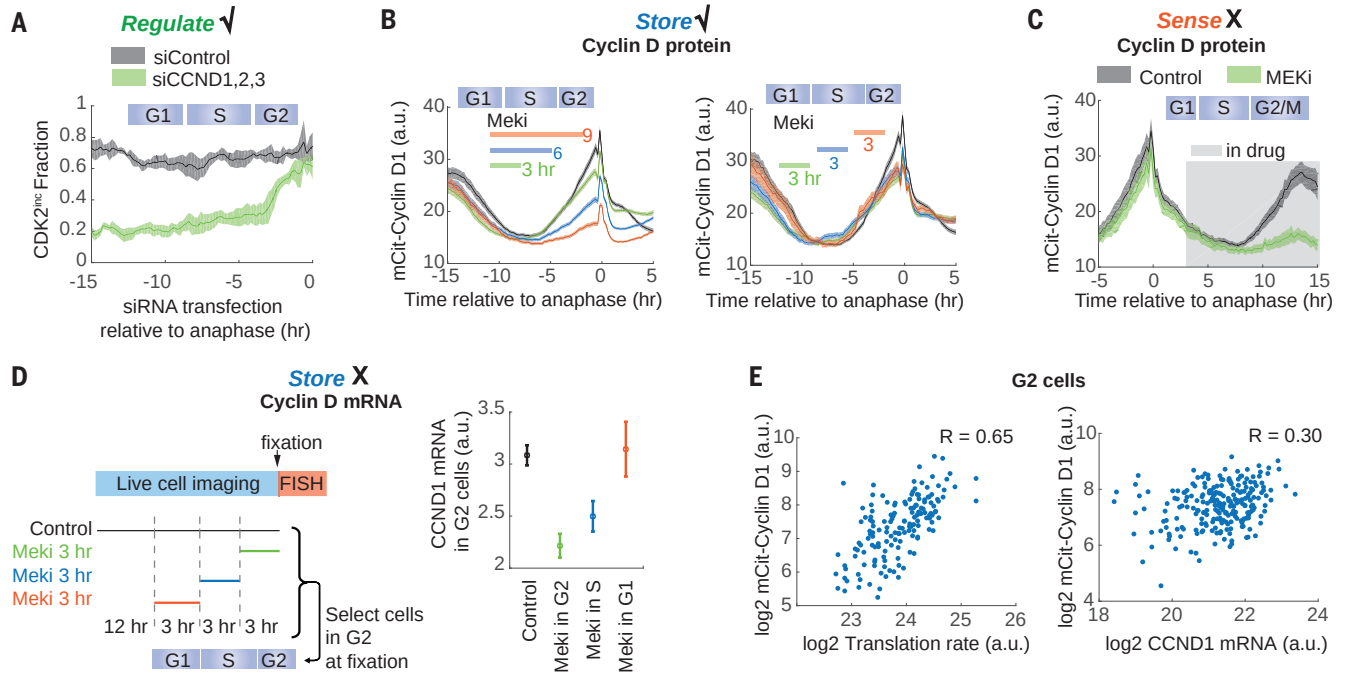


Fig. 2. Cyclin D acts as a downstream effector of the MAPK integrator. (**A**) Cyclin D regulates cell proliferation. The fraction of $CDK2^{inc}$ daughter cells is plotted against the time of small interfering RNA (siRNA) addition. siControl, nontargeting siRNA; siCCND1,2,3, siRNAs against CCND1, CCND2, and CCND3. (**B**) Cyclin D1 protein levels in G_2 correlate negatively with the duration of the Meki treatment in the ongoing cell cycle, regardless of the cell cycle phase of the treatment. Cells expressing mCitrine-cyclin D1 from the endogenous locus were treated for 0 (black curve), 3, 6, or 9 hours (left), or for 3 hours (right) with Meki during imaging; the timing of treatment is indicated by colored bars. a.u., arbitrary units. (**C**) Cyclin D1 protein levels in G_1 do not sense Meki treatment. The experimental setup is the same as in (B), except that Meki was left in once added. (**D**) Meki treatment in G_1 does not reduce cyclin D1 mRNA levels in G_2. Time-lapse imaging of CDK2 activity in asynchronous cells was followed by a 3-hour Meki treatment 0, 3, or 6 hours before fixation. Correspondingly, G_2 cells at the time of fixation (10 to 12 hours after anaphase) received Meki treatment in the G_2, S, or G_1 phase, respectively. Cyclin D1 mRNA levels were measured by RNA fluorescence in situ hybridization (FISH) as the total signal in individual G_2 cells. (**E**) Cyclin D1 protein levels strongly correlate with protein translation rate in G_2 cells. Single-cell translation rates were measured with the O-propargyl-puromycin (OPP) assay, where cells were pulsed with OPP for 24 min before fixation (*26*) and quantified as the integrated intensity of incorporated OPP within each cell. G_2 cells were identified, and cyclin D1 mRNA levels were measured as in (D). Each dot represents a single cell. R value is calculated as the Pearson correlation coefficient. mCit, mCitrine. With the exception of (E), all data are plotted as means ± 95% confidence intervals shown as shaded bands [(A), (B), and (C)] or error bars (D). All data are from MCF10A cells.

defect on overexpression (figs. S6 and S7). Moreover, cells receiving Meki for 3 hours in G_1 alone do not show reduced cyclin D1 mRNA in G_2 (Fig. 2D), which suggests that cyclin D1 mRNA levels, in contrast to cyclin D1 protein levels, do not store a cell's MAPK history. This discrepancy points to a post-transcriptional regulation step of cyclin D in relaying MAPK history to cyclin D protein levels in G_2.

Cyclin D is an unstable protein and thus is acutely sensitive to changes in translation rate (fig. S8). Its yeast homolog, Cln3, has been proposed to act as a translational sizer, the abundance of which reflects the translation rate of cells (*23*, *24*). The level of cyclin D protein strongly correlates with the translation rate in individual G_2 cells, much more so than it correlates with its mRNA levels (Fig. 2E). The building of protein mass is naturally an integration process. We therefore tested whether protein synthesis might act as the integrator that reflects the history of MAPK signaling.

Direct inhibition of translation with puromycin or cycloheximide impairs cell cycle progression and blocks mother cell mitoses and therefore cannot be used to study daughter cell proliferation (fig. S9, A and B). We circumvented this problem by indirectly reducing translation in mother cells using the mTor (mammalian target of rapamycin) inhibitor torin, which only mildly extends the length of the mother cell cycle (fig. S9, C and D). Torin treatment impairs proliferation of daughter cells in a duration-dependent, cell cycle phase–independent manner (Fig. 3A), which suggests that unperturbed translation throughout the mother cell cycle is required for maximal daughter cell proliferation [criterion (i)]. Translation rates rapidly decrease after a short Meki treatment in a cell cycle phase–independent manner, which suggests that translation rates sense Mek activity throughout the cell cycle [criterion (ii); Fig. 3, B and C]. Longer Meki treatment leads to further decrease of translation (Fig. 3, B and C), a duration-dependent effect similar to that observed on cyclin D1 levels in G_2 and on the fraction of proliferative daughter cells (Fig. 3D). Cells receiving a brief Meki treatment in the G_1 or S phase maintain lower translation rates in G_2, which suggests that protein translation stores the history of MAPK activity [criterion (iii); Fig. 3E] and can function as an integrator of mitogen signaling.

To determine whether protein translation is the integrator that regulates the cyclin D rise in G_2 and consequent proliferation decisions in daughter cells, we sought to test whether enhanced translation could rescue the Meki-induced proliferation defect. Inspired by the observation that CDK4/6 inhibition is the only condition out of many tested where cells accumulate protein mass in the absence of cell cycle progression (*25*), we pretreated cells with palbociclib for 24 hours to increase their mass (and hence ribosome number) (Fig. 3F and fig. S9E). We then released these cells into the cell cycle; treated them with Meki in the G_1 or S phase for 1, 3, or 6 hours; and tracked daughter

cell proliferation (Fig. 3F). Pretreating cells with CDK4/6 inhibitor before Meki treatment restores cyclin D1 levels in G_2 and proliferation in daughter cells (Fig. 3, F and G). Therefore, we conclude that the history of MAPK activity is stored in the translation rate, which in turn influences the proliferation of daughter cells.

Contrary to the long-standing paradigm that cells evaluate their mitogenic environment in a window before the R point in G_1, our data indicate that cells integrate mitogen signaling throughout the entire mother cell cycle to modulate the proliferation-quiescence decision in daughter cells (Fig. 4A). This proposed model can be illustrated by a simple mechanical analogy, where water constantly drips into a bucket, and it is the integral of the dripping—the

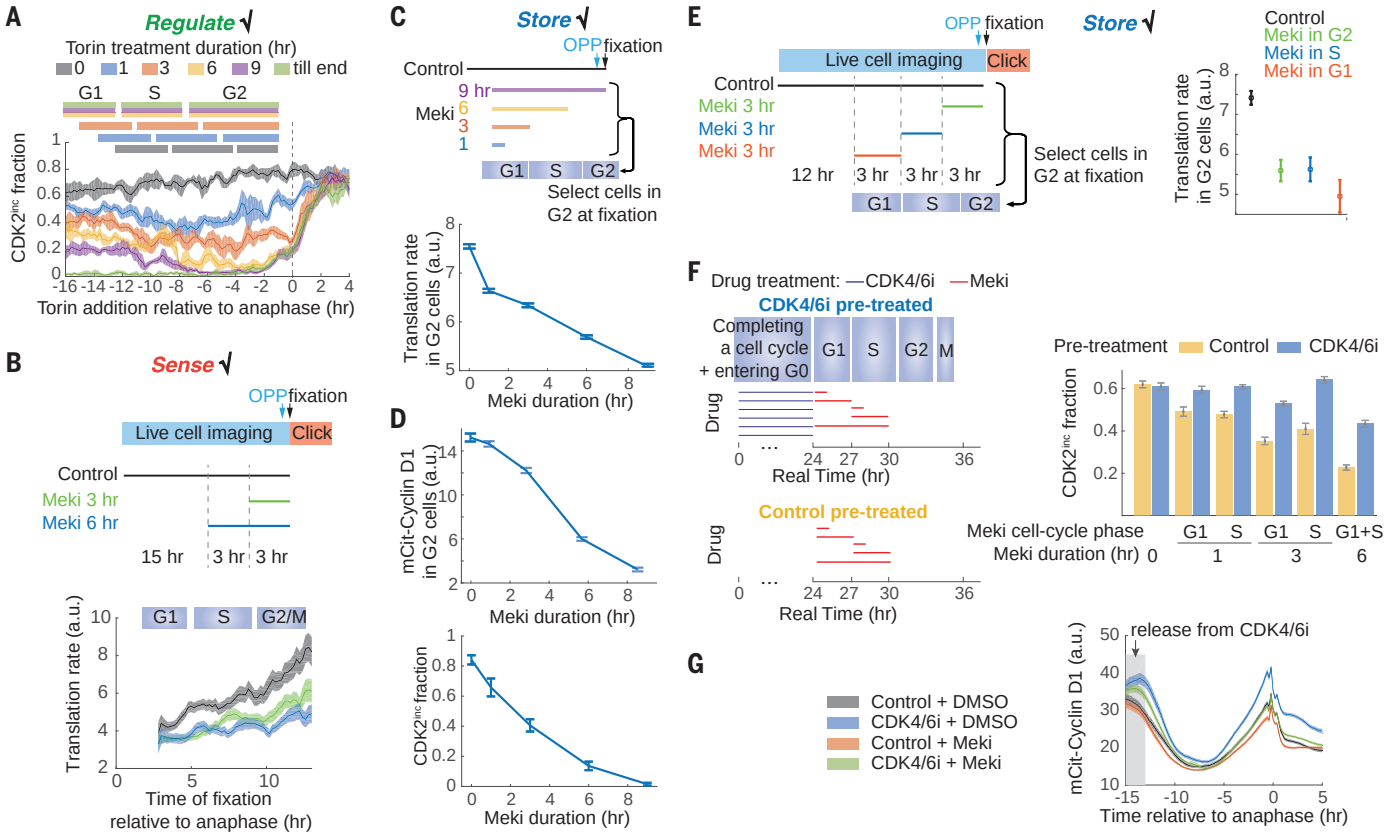
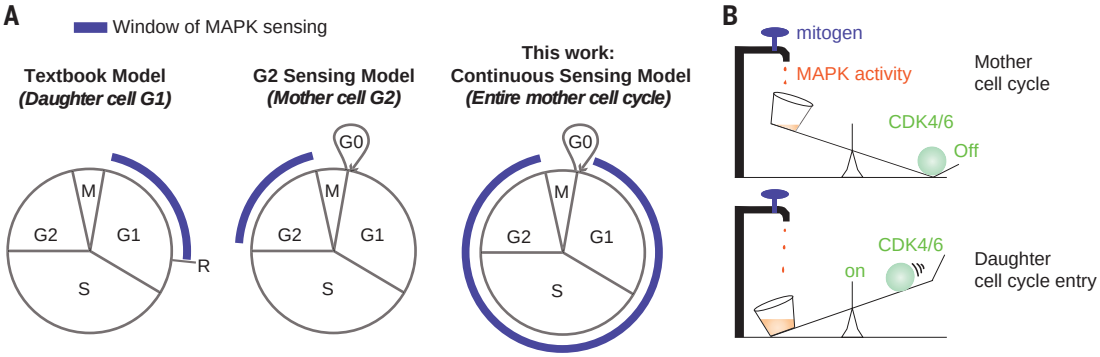


Fig. 3. Global protein translation rate integrates MAPK activity. (**A**) Perturbation of protein translation by torin in the mother cell cycle impairs daughter cell cycle entry in a duration-dependent manner. The experimental setup is the same as in Fig. 1E except that torin is used. Cell cycle phases are derived from fig. S9, C and D. (**B**) Translation rate can rapidly sense MAPK activity throughout the cell cycle. Time-lapse imaging of CDK2 activity in asynchronous cells was followed by an OPP assay, as in Fig. 2E. The OPP signal was then reconstructed as a function of the time since anaphase. (**C**) Meki reduces the G_2 phase translation rate in a duration-dependent manner. (**D**) Duration-dependent effect of Meki on the G_2 phase cyclin D1 protein levels and on the fraction of $CDK2^{inc}$ daughter cells. (**E**) Translation rates can store past MAPK activity. The experimental setup is the same as in (B) except that cells were fixed at 0, 3, or 6 hours after the 3-hour Meki treatment. Correspondingly, G_2 cells at the time of fixation (10 to 12 hours after anaphase) received Meki treatment in the G_2, S, or G_1 phase, respectively. (**F**) Enhancing translation by pre-enlarging cells using CDK4/6i can rescue the Meki-induced proliferation defect. (**G**) CDK4/6i pretreatment can rescue the Meki-induced reduction of cyclin D1 levels. All data are from MCF10A cells and are plotted as means ± 95% confidence intervals, shown as shaded bands [(A), (B), and (G)] or error bars [(C), (D), (E), and (F)].

Fig. 4. Models. (**A**) An updated model of mitogen sensing. R, restriction point. (**B**) A mechanical metaphor of how MAPK signaling promotes cell cycle entry.



weight of the bucket—that eventually flips the proliferation-quiescence switch (Fig. 4B). We propose that protein synthesis functions as the bucket to record the history of mitogenic signals throughout the mother cell cycle. Given that translation rate is strongly correlated with cell size and cell growth, cell growth itself may be the bucket (the integrator); however, it is not currently possible to measure single-cell growth at the precision required to test this idea. The MAPK history is decoded into cyclin D protein levels in the mother cell G_2 to regulate the proliferation of daughter cells. This ensures that cells achieve a threshold protein synthesis rate before committing to the cell cycle. Notably, by virtue of the temporal integration, this system constitutes a form of cellular memory, in which the past experience of mitogenic signals during the entire mother cell cycle influences the fraction of proliferating daughter cells.

REFERENCES AND NOTES

1. S. Meloche, J. Pouysségur, *Oncogene* **26**, 3227–3239 (2007).
2. J. E. Toettcher, O. D. Weiner, W. A. Lim, *Cell* **155**, 1422–1434 (2013).
3. D. O. Morgan, *The Cell Cycle, Principles of Control* (New Science Press, 2007).
4. A. B. Pardee, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 1286–1290 (1974).
5. R. A. Weinberg, *Cell* **81**, 323–330 (1995).
6. S. L. Spencer *et al.*, *Cell* **155**, 369–383 (2013).
7. M. Hitomi, D. W. Stacey, *Mol. Cell. Biol.* **19**, 4623–4632 (1999).
8. H. W. Yang, M. Chung, T. Kudo, T. Meyer, *Nature* **549**, 404–408 (2017).
9. J. Moser, I. Miller, D. Carter, S. L. Spencer, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E8219–E8227 (2018).
10. S. Gookin *et al.*, *PLOS Biol.* **15**, e2003268 (2017).
11. J. G. Albeck, G. B. Mills, J. S. Brugge, *Mol. Cell* **49**, 249–261 (2013).
12. T. E. Gillies, M. Pargett, M. Minguet, A. E. Davies, J. G. Albeck, *Cell Syst.* **5**, 549–563.e5 (2017).
13. J. J. Liu *et al.*, *Mol. Cell. Biol.* **15**, 3654–3663 (1995).
14. J. N. Lavoie, G. L'Allemain, A. Brunet, R. Müller, J. Pouysségur, *J. Biol. Chem.* **271**, 20608–20616 (1996).
15. P. W. Hinds *et al.*, *Cell* **70**, 993–1006 (1992).
16. S. A. Ezhevsky *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10699–10704 (1997).
17. C. J. Sherr, J. M. Roberts, *Genes Dev.* **9**, 1149–1163 (1995).
18. M. Hitomi, D. W. Stacey, *Curr. Biol.* **9**, 1075–1084 (1999).
19. T. Zerjatke *et al.*, *Cell Rep.* **19**, 1953–1966 (2017).
20. C. Albanese *et al.*, *J. Biol. Chem.* **270**, 23589–23597 (1995).
21. P. M. Burch, Z. Yuan, A. Loonen, N. H. Heintz, *Mol. Cell. Biol.* **24**, 4696–4709 (2004).
22. J. I. Daksis, R. Y. Lu, L. M. Facchini, W. W. Marhin, L. J. Penn, *Oncogene* **9**, 3635–3645 (1994).
23. M. Polymenis, E. V. Schmidt, *Genes Dev.* **11**, 2522–2531 (1997).
24. A. Litsios *et al.*, *Nat. Cell Biol.* **21**, 1382–1392 (2019).
25. M. B. Ginzberg *et al.*, *eLife* **7**, e26957 (2018).
26. J. Liu, Y. Xu, D. Stoleru, A. Salic, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 413–418 (2012).

ACKNOWLEDGMENTS

We thank J. Mansfeld for RPE-hTERT Flp-In parental and RPE-hTERT cyclin D1-mVenus cell lines; J. Albeck for MCF10A Erk-sensor cell line; N. Ahn and J. Albeck for insightful discussions and comments on the manuscript; BioFrontiers Computing Core for computing support; and the Spencer laboratory for general discussion and assistance. **Funding:** This work was supported primarily by an NIH K22 Early-Career Investigator award (1K22CA188144-01) and an NIH Director's New Innovator award (DP2-CA238330), as well as an American Cancer Society Research Scholar grant (RSG-18-008-01), a Beckman Young Investigator award, a Boettcher Webb-Waring Early-Career Investigator award, a Kimmel Scholar award (SKF16-126), a Pew-Stewart Scholar award, and a Searle Scholar award (SSP-2016-1533) to S.L.S. **Author contributions:** M.M. and S.L.S. designed the study, interpreted the data, and wrote the manuscript. M.M. performed most of the experiments and data analysis with help from Y.R. C.T. developed the tracking code for RPE (retinal pigment epithelial) cells. **Competing interests:** The authors declare no conflict of interest. **Data and materials availability:** All numerical data are available in the BioStudies database under accession number S-BSST314. Materials are available by contacting S.L.S.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1261/suppl/DC1
Materials and Methods
Figs. S1 to S9
References (*27*–*36*)
MDAR Reproducibility Checklist

22 July 2019; resubmitted 19 December 2019
Accepted 17 March 2020
Published online 2 April 2020
10.1126/science.aay8241

BIOTECHNOLOGY

Very fast CRISPR on demand

Yang Liu[1]*, Roger S. Zou[2]*, Shuaixin He[1], Yuta Nihongaki[3], Xiaoguang Li[3], Shiva Razavi[2]†, Bin Wu[1,5,6]‡, Taekjip Ha[1,2,4,7]‡

CRISPR-Cas systems provide versatile tools for programmable genome editing. Here, we developed a caged RNA strategy that allows Cas9 to bind DNA but not cleave until light-induced activation. This approach, referred to as very fast CRISPR (vfCRISPR), creates double-strand breaks (DSBs) at the submicrometer and second scales. Synchronized cleavage improved kinetic analysis of DNA repair, revealing that cells respond to Cas9-induced DSBs within minutes and can retain MRE11 after DNA ligation. Phosphorylation of H2AX after DNA damage propagated more than 100 kilobases per minute, reaching up to 30 megabases. Using single-cell fluorescence imaging, we characterized multiple cycles of 53BP1 repair foci formation and dissolution, with the first cycle taking longer than subsequent cycles and its duration modulated by inhibition of repair. Imaging-guided subcellular Cas9 activation further facilitated genomic manipulation with single-allele resolution. vfCRISPR enables DNA-repair studies at high resolution in space, time, and genomic coordinates.

RNA-guided DNA targeting with CRISPR-Cas9 has revolutionized biomedical research for genome editing and beyond (*1*). After genomic DNA cleavage by Cas9, DNA damage response (DDR) proteins are recruited to initiate complex repair processes (*2*). Although DDR is known to be influenced by factors such as target sequence (*3*, *4*), cell cycle (*5*), and chromatin dynamics (*6*), the precise timing and sequence of cellular events require further investigation. Cas9 has potential as a tool to study the dynamics of DDR but currently lacks the necessary level of control to initiate precise DNA damage on demand. To unveil the sequence of Cas9-induced DDR events in living cells, an inducible Cas9 system with the spatiotemporal resolution that matches the rapidity and subcellularity of DDR would be powerful.

Numerous inducible Cas9 systems have been developed (*7*–*11*). However, these methods often exhibit compromised function in the engineered proteins, coarse temporal control in the hour time scale (because Cas9 still has to find the target after induction), and no spatial control or control at the millimeter length scale at best.

Here, we report a very fast CRISPR-Cas9 system (vfCRISPR) that allows genome editing on demand at the submicrometer space scale and the second time scale. Through synchronized double-strand break (DSB) induction followed by complementary biochemical, sequencing, and imaging-based assays, we characterized the early molecular events that underlie the initiation and progression of DNA repair with high spatiotemporal precision.

The design principle of vfCRISPR is based on the *Streptococcus pyogenes* Cas9 (Cas9 henceforth) cleavage mechanism. The protospacer adjacent motif (PAM)–proximal 9- to 10-bp region of guide RNA (gRNA) governs Cas9 binding to its target DNA, whereas additional base pairing at the PAM-distal region (10 to 20 bp) is required for cleavage (*12*, *13*). Mismatches in the PAM-distal region prevent full unwinding of target DNA (*14*) and

[1]Department of Biophysics and Biophysical Chemistry, Johns Hopkins University, Baltimore, MD, USA. [2]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA. [3]Department of Cell Biology, Johns Hopkins University, Baltimore, MD, USA. [4]Department of Biophysics, Johns Hopkins University, Baltimore, MD, USA. [5]The Solomon H. Snyder Department of Neuroscience, Johns Hopkins University, Baltimore, MD, USA. [6]The Center for Cell Dynamics, Johns Hopkins University, Baltimore, MD, USA. [7]Howard Hughes Medical Institute, Baltimore, MD, USA.
*These authors contributed equally to this work.
†Present address: Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA.
‡Corresponding author. Email: tjha@jhu.edu (T.H.); bwu20@jhu.edu (B.W.)

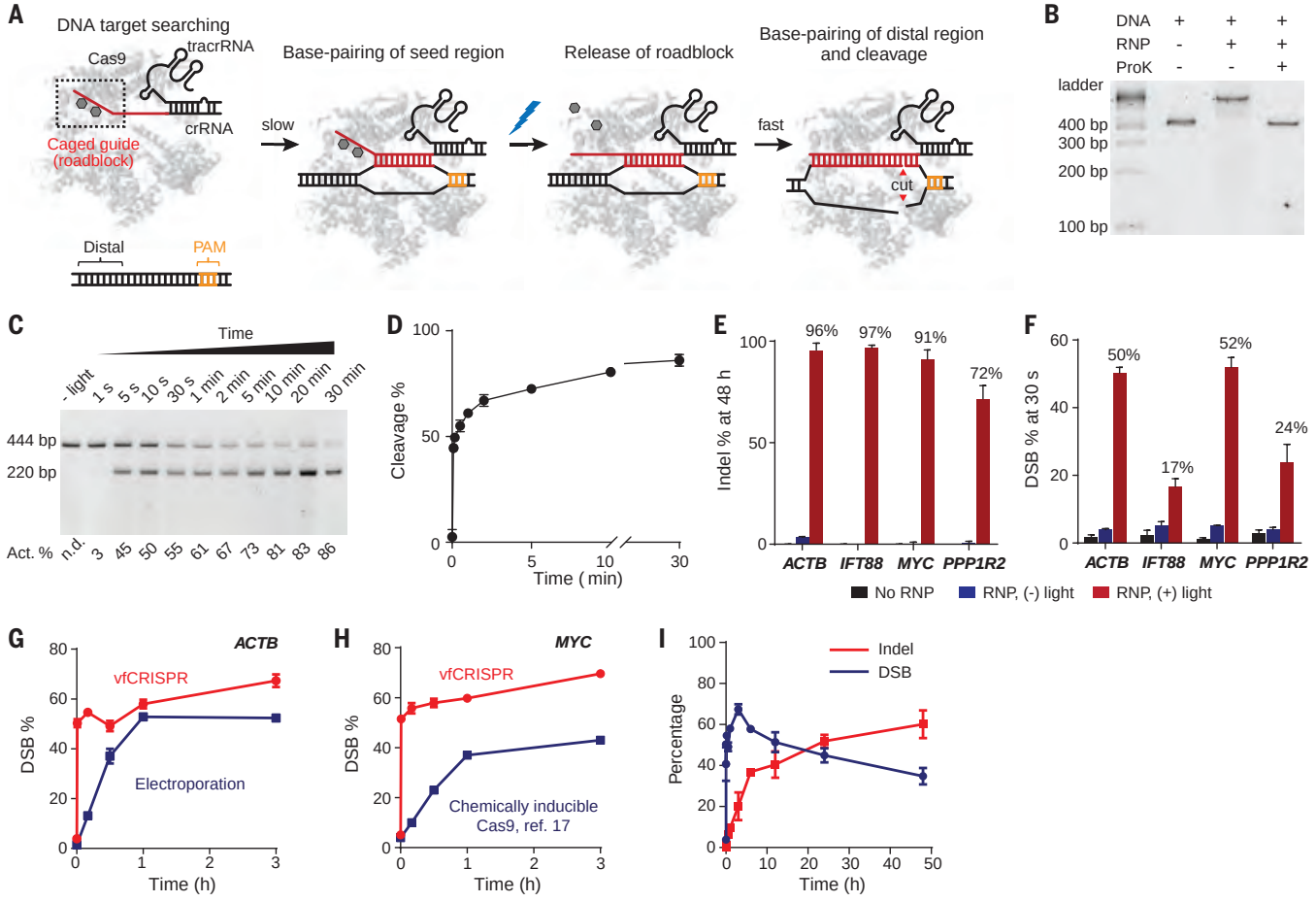


Fig. 1. Characterization of vfCRISPR in vitro and in cells. (**A**) Schematic of Cas9 activation by modulating base pairing between the PAM-distal region of cgRNA and genomic DNA. (**B**) Without light, Cas9/cgRNA ribonucleoprotein (RNP) bound to target DNA without cleavage, causing a clear band shift. Proteinase K degraded Cas9, causing target DNA to shift back to the original position. (**C** and **D**) Fast and efficient in vitro cleavage kinetics of Cas9 after light activation. (**E**) Indels detected by high-throughput sequencing of PCR-amplified genomic DNA extracted from cells without RNP, with RNP but no light, and with RNP 48 hours after light activation. (**F**) DSBs detected by DSB-droplet digital PCR of genomic DNA extracted from cells without RNP, with RNP but no light, and with RNP 30 s after light activation. (**G** and **H**) Percentage of DSBs over time using vfCRISPR (red) compared with either RNP electroporation [(G), target sequence at *ACTB*] or a chemically inducible system [(H), target sequence at *MYC*]. (**I**) DSBs and normalized indels (see the materials and methods) at *ACTB* over time after Cas9 activation.

conformational changes of the HNH domain (*15*) required for cleavage. On the basis of this mechanistic understanding, we replaced two or three uracils at the PAM-distal region of crRNA with light-sensitive, 6-nitropiperonyloxymethyl–modified deoxynucleotide thymine caged nucleotides (*16*), forming a caged gRNA (cgRNA) when hybridized to wild-type transactivating CRISPR RNA (tracrRNA) (Fig. 1A). The Cas9/cgRNA complex retains the ability to bind its target DNA but cannot cleave because the steric hindrance imposed by the caging groups prevents full DNA unwinding and nuclease activation. Upon light stimulation at 365 or 405 nm, the caging groups are removed and the prebound, now-activated Cas9/cgRNA complex rapidly cleaves target DNA.

An electrophoretic mobility shift assay confirmed that Cas9/cgRNA stably bound to target DNA without light and no cleavage was observed (Fig. 1B). After uncaging with light, the Cas9/cgRNA complex efficiently cleaved DNA within seconds in vitro (Fig. 1, C and D, and fig. S1).

Next, we characterized the activity of vfCRISPR in human embryonic kidney 293 cells by targeting four endogenous loci and found light-induced indel efficiency up to 97%, whereas cells without light exposure had almost no detectable indels (Fig. 1E and fig. S2). Cells exposed to this dosage of light exhibited no apparent phototoxicity (fig. S3). Approximately 50% of DNA cleavage was found within 30 s after light activation (Fig. 1F and fig. S4). Compared with other Cas9 induction methods, vfCRISPR exhibited much faster cleavage kinetics and higher cleavage efficiency (*17*, *18*) (Fig. 1, G and H). We attribute the very fast kinetics to skipped nuclear localization or target-searching steps, and the higher cleavage efficiency to the use of wild-type Cas9. Genome-wide analysis of off-target editing using GUIDE-seq (*19*) also revealed reduced off-target activity compared with wild-type gRNA (fig. S5), consistent with improved specificity from deoxyribonucleotide incorporation into the guide RNA (*20*). These experiments demonstrated that cgRNA enables very fast and efficient inducible DNA cleavage in mammalian cells.

With a precisely defined time for cleavage, vfCRISPR allowed us to investigate the generation and repair kinetics of Cas9-mediated DSBs. We measured the percentage of DSBs and indels as a function of time after Cas9 activation at multiple target sites and adopted mathematical models to describe the kinetics of DSB and indel formation (*21*) (Fig. 1I, fig. S6, and supplementary materials, models I and II). Model fitting led us to hypothesize the recutting of +1 insertion DNA at *ACTB*, which

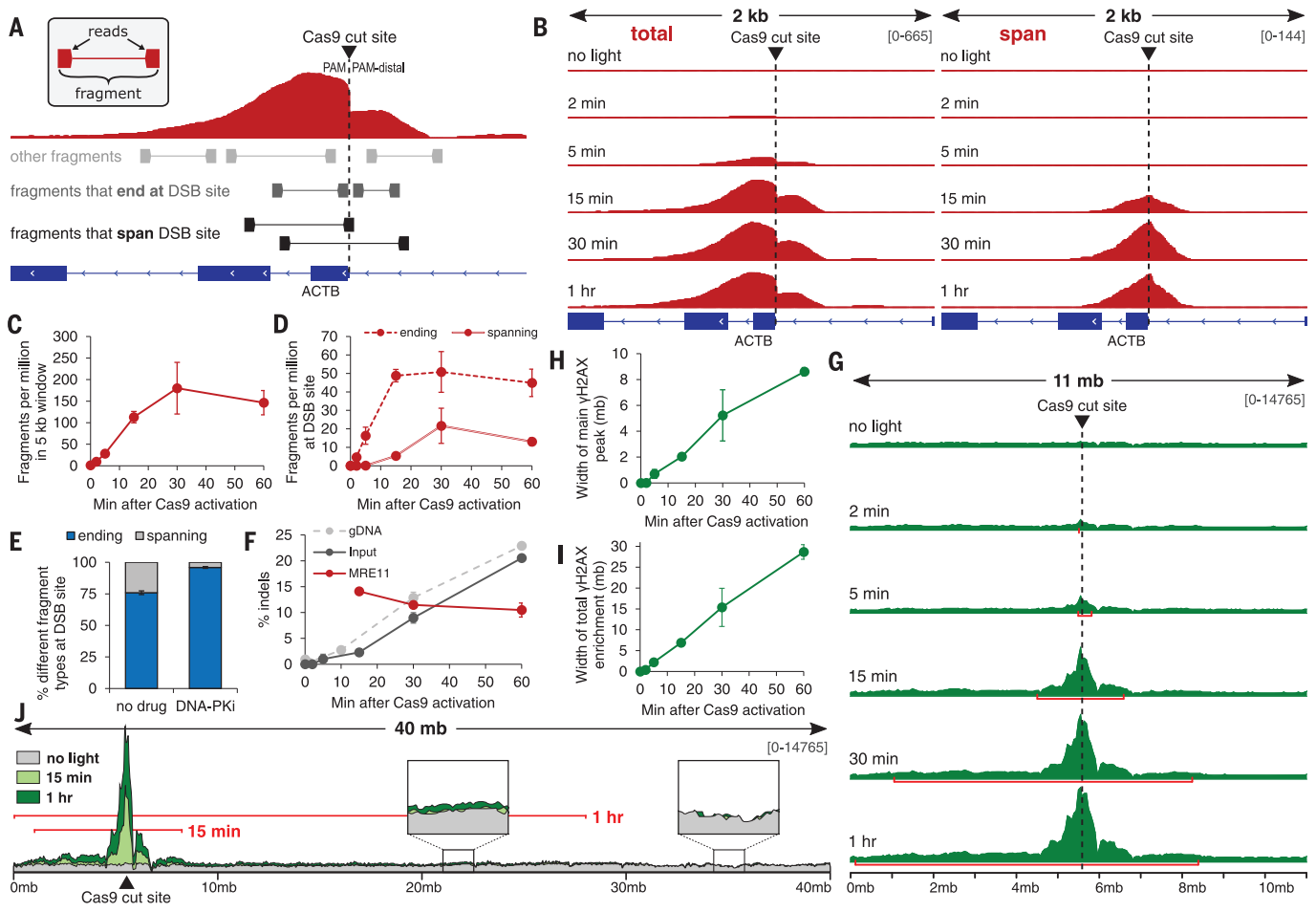


Fig. 2. trChIP-seq reveals dynamics of MRE11 recruitment and H2AX phosphorylation after synchronized Cas9-induced DSBs. (**A**) Schematic of MRE11 ChIP-seq analysis for paired-end reads. (**B**) Visualization of MRE11 peak features over time after Cas9 activation. The left column ("total") piles all fragments. The right column ("span") only piles fragments that span the cleavage site. (**C**) Fragments per million in the 5-kb window around the cleavage site over time. (**D**) Fragments per million that either span or start and/or end at the cleavage site over time. Fragments that start and/or end at the DSB site are enriched first, followed by an ~15-min delay by fragments that span the DSB site. (**E**) Proportion of fragments that span the DSB site that are depleted with DNA-PKcs inhibition. (**F**) Percentage of indels calculated from deep sequencing of PCR amplicons from both ChIP input (dark gray) and MRE11 ChIP (red) DNA that span the cleavage site. Indel kinetics from Fig. 1I are included for comparison (light gray, dashed, "gDNA"). (**G**) γH2AX enrichment over time in an 11-mb window around the cleavage site. Red bars mark the width of the "main" γH2AX peak detected using MACS2. (**H**) Width of the main γH2AX peak over time detected using MACS2. (**I**) Width of total enrichment detected using Student's *t* test with Bonferroni correction ($P < 0.05$) comparing no-light results with all after-activation results. (**J**) Illustration of enrichment up to ~30 mb along Chr7:1-40,000,000 ($P < 0.05$). Red bars mark the width of total γH2AX enrichment detected using Student's *t* test.

we subsequently verified both in vitro and in cells through recutting of a monoclonal cell line with a pure +A indel product at *ACTB* (figs. S7 and S8).

Using highly synchronized DNA cleavage, we performed time-resolved chromatin immunoprecipitation followed by sequencing (trChIP-seq) to track the recruitment of MRE11, which forms the MRN complex with Rad50 and Nbs1, to the *ACTB* cleavage site (*18*) (Fig. 2A). We observed rapid MRE11 recruitment that reached half-maximal signal between 5 and 15 min (Fig. 2, B and C, and fig. S9A). This is slower than recruitment of another MRN component, Rad50, after laser microirradiation (*22*), potentially because of the delay in exposure of Cas9-induced DSBs (fig. S9B). Both ChIP-seq and probe-based ChIP-quantitative polymerase chain reaction (qPCR) detected the emergence of MRE11-bound DNA that spanned the cleavage site 15 min after Cas9 activation (Fig. 2, B and D, and fig. S10A). This spanning population, which we attribute to repaired DNA still bound by MRE11, was present across different target sequences and cell types (*18*) (fig. S10, B to D, and fig. S11). Inhibition of the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs) with KU-0060648 led to a reduction in spanning fragments with a concomitant increase in fragments that ended at the cut site (Fig. 2E and fig. S12). Deep amplicon sequencing of spanning DNA revealed consistent 10 to 15% indels, whereas total indels rose from 2 to 20% within that period of time (Fig. 2F and fig. S13). Together, these results are consistent with transient MRE11 retention on ligated genomic DNA that was processed in a DNA-PKcs–dependent manner.

H2AX is known to be phosphorylated (γH2AX) for 1 to 2 Mb around a DSB undergoing active repair (*23*), but the dynamics of initial γH2AX spreading is unknown. trChIP-seq for γH2AX revealed rapid expansion of a "main peak" around the cleavage site at the speed of ~150 kb/min, reaching 8 Mb at 1 hour (Fig. 2, G and H, and fig. S14). Statistical testing (see the materials and methods) revealed another layer of γH2AX enrichment

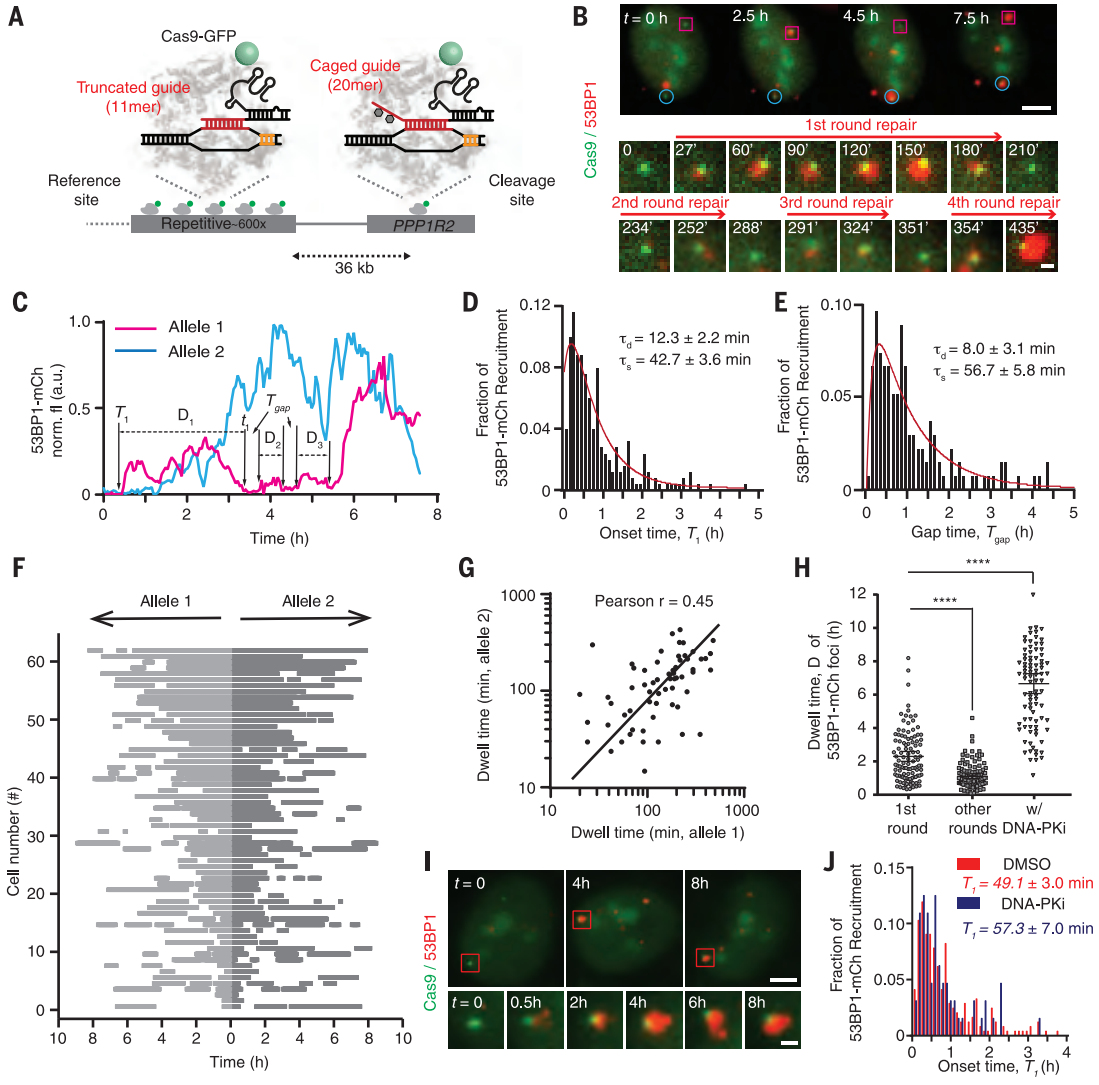


Fig. 3. Live-cell imaging reveals the spatiotemporal dynamics of 53BP1 after synchronized Cas9 DSBs. (**A**) Schematic of orthogonal genomic imaging and cleavage using Cas9-EGFP. (**B**) Snapshots showing 53BP1 recruitment to two different *PPP1R2* alleles (magenta square and cyan circle). Scale bar, 5 μm. Magnified images show multiple cycles of 53BP1 foci at one *PPP1R2* allele (magenta square). Scale bar, 0.7 μm. (**C**) Fluorescence intensity traces of 53BP1-mCherry for both alleles monitored in (B). T_1, t_1, and D_1 denote the start, end, and dwell times for the first cycle of DNA repair, respectively. The time interval T_{gap} is calculated as $T_{n+1} - t_n$ (e.g., $T_2 - t_1$). (**D**) Histogram of initial 53BP1 recruitment time (T_1) at the *PPP1R2* locus after light stimulation. (**E**) Histogram of time interval (T_{gap}) between consecutive 53BP1 cycles for 8 hours. (**F**) Rastergram of 53BP1 foci at 124 paired alleles in 62 cells. Each row displays time courses of 53BP1 foci at a pair of alleles residing in the same nucleus. Gray bars indicate presence of 53BP1-mCherry at each *PPP1R2* allele. Cells are ranked by the mean dwell time of the first 53BP1 recruitment at two alleles (longest to shortest). (**G**) Scatter plot showing positive correlation in dwell time between two alleles in the same cell nucleus ($r = 0.45$). (**H**) Dwell time of 53BP1 foci at *PPP1R2* for the first 53BP1 cycle ($n = 167$ foci in 5 biological replicates), later cycles ($n = 109$ foci in 5 biological replicates), or after DNA-PKcs inhibition (DNA-PKi; $n = 92$ foci in 3 biological replicates). Unpaired t test was performed (****$P < 0.0001$). Error bars indicate 95% confidence interval. (**I**) Snapshots showing longer 53BP1 dwell time in cells with DNA-PKcs inhibition. Scale bars, 5 and 1 μm. (**J**) 53BP1 recruitment time was unchanged ($P > 0.2$) in DNA-PKcs–inhibited cells.

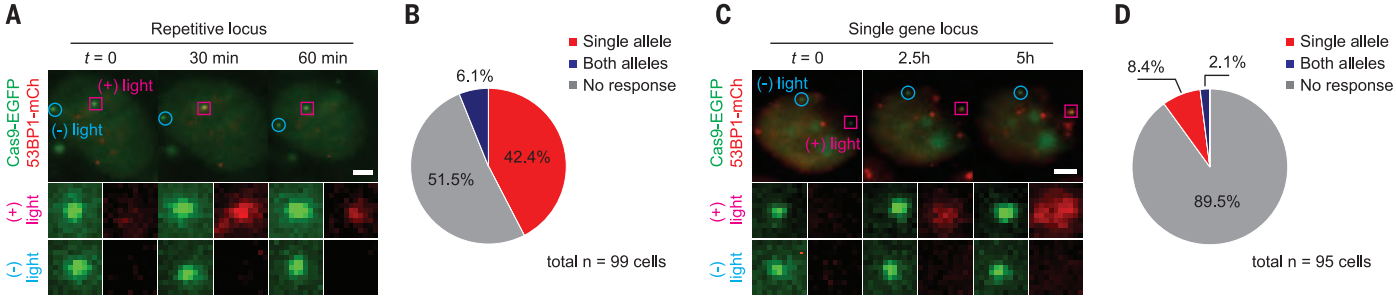


Fig. 4. DSB cleavage at single-allele resolution using subcellular Cas9 activation. (**A**) Single-cell snapshots showing targeted cleavage of Ch3Rep at one of the two alleles (pink). Scale bar, 5 μm. (**B**) Summary of 53BP1-mCh recruitment to Ch3Rep alleles. "Single allele" (red) indicates 53BP1-mCherry recruitment to the targeted allele only; "both alleles" (blue) indicates recruitment to both the targeted and nontargeted alleles; "no response" (gray) indicates no recruitment to either allele ($n = 99$ foci in 4 biological replicates). (**C**) Single-cell snapshots showing targeted cleavage of a single *PPP1R2* allele (pink). Scale bar, 5 μm. (**D**) Summary of 53BP1-mCh recruitment to *PPP1R2* alleles ($n = 95$ foci in 4 biological replicates). Single-allele specificity was calculated by dividing the percentage of monoallelic activation by the percentage of total activation (both monoallelic and biallelic activation). For repetitive cutting (B), this was 42.4/(42.4 + 6.1) × 100 = 87.4%, whereas for single cutting (D), it was 8.4/(8.4 + 2.1) × 100 = 80%.

that expanded linearly at 460 kb/min and spanned up to 30 Mb at 1 hour (Fig. 2, I and J, and fig. S15). To our knowledge, γH2AX enrichment up to tens of megabases has not been previously reported. Although this may be a feature specific to Cas9-induced DSBs, synchronized, high-efficiency cleavage with vfCRISPR may have contributed to the detection of lower levels of enrichment.

Next, we performed single-cell fluorescence imaging to capture the dynamics of repair protein recruitment induced by vfCRISPR. To a monoclonal U-2 OS cell line stably expressing Cas9-EGFP, we cotransfected a truncated gRNA (11-mer in the protospacer) targeting a highly repetitive region in chromosome 3 (Ch3Rep) with a cgRNA targeting the *PPP1R2* gene, only 36 kb downstream of Ch3Rep (*24*–*26*). Upon light activation, a single DSB is generated at *PPP1R2*, which is fluorescently marked by an array of Cas9-EGFPs decorating Ch3Rep (Fig. 3A). We confirmed recruitment of multiple endogenous repair factors (pATM, MDC1, 53BP1, and γH2AX) to the single break sites (fig. S16, A to D). γH2AX foci size increased over time (fig. S16D), consistent with the γH2AX spreading reported by trChIP-seq.

To track real-time repair dynamics through live-cell imaging, we stably coexpressed 53BP1-mCherry (*27*) with Cas9-EGFP in U-2 OS cells (Fig. 3, A to C). Most of the Cas9-EGFP–labeled alleles colocalized with 53BP1-mCherry foci over the course of 8 hours after Cas9 activation (fig. S17A and movie S1), indicating efficient cleavage at *PPP1R2*. The onset time for 53BP1 recruitment (T_1) was heterogeneous between cells and alleles, with most foci appearing within 1 hour (Fig. 3D). Maximum likelihood estimation of a two-step mathematical model for T_1 yielded a DSB detection time (τ_d) of 12 ± 2.2 min and a Cas9 target-searching time (τ_s) of 43 ± 3.6 min (Fig. 4C and supplementary materials).

Each 53BP1 focus underwent a cycle of enlargement and dissolution, with most exhibiting more than one and up to five cycles over 8 hours, consistent with a previous report (*26*) (fig. S17B). The estimated τ_s and τ_d for subsequent rounds of 53BP1 recruitment agreed well with those for the first round, suggesting that each 53BP1 cycle corresponds to at least one repair event (Fig. 3E and table S6).

Duration of the initial 53BP1 cycle (D_1) was on average significantly longer than subsequent cycles (Fig. 3H). D_1 varied over a wide range and was positively correlated between two alleles in the same cell but not between different cells (Fig. 3, F and G, and figs. S18 and S19), suggesting that stochastic differences in chromatin environments between the two alleles is not the main reason for the large variation in D_1. Inhibition of ATM using KU-0055933 eliminated 53BP1 foci (fig. S20), consistent with its role as an upstream regulator of 53BP1 recruitment. Inhibition of DNA-PKcs prolonged 53BP1 foci without affecting 53BP1 recruitment (Fig. 3, H to J, and fig. S20), leading to fewer 53BP1 cycles and further supporting our interpretation that 53BP1 cycles mark successive rounds of DSBs and repair (figs. S21 to S23).

Finally, we extended vfCRISPR to spatially manipulate single genomic alleles. Both Ch3Rep alleles were bound by an array of Cas9/cgRNAs within the same nucleus. We focused a 405-nm laser beam to one Ch3Rep allele, which locally activated Cas9 to cleave the targeted allele while keeping the other one intact (fig. S24). Almost half of cells exhibited 53BP1 recruitment to the targeted allele within 1 hour, whereas only 6% of cells showed recruitment to both alleles, showing single-allele specificity (Fig. 4, A and B, and movie S2). We also demonstrated allele-specific manipulation of a nonrepetitive cleavage site at *PPP1R2* (Fig. 4, C and D, and movie S3). We observed only one round of 53BP1-mCherry recruitment, likely because only one round of cleavage was possible owing to the activation of only a small subset of Cas9/cgRNA in the nucleus (fig. S25). Laser beam alone was not responsible for 53BP1 recruitment (fig. S26), and conditions were optimized to maximize single-allele specificity, which was 80% or higher for both experimental schemes (Fig. 4, B and D). The capability of vfCRISPR to manipulate single genomic alleles with high specificity motivates applications such as simplifying generation of heterozygous mutants (*28*) and potentially reducing and/or eliminating off-target genome editing.

To the best of our knowledge, vfCRISPR provides the highest spatial and temporal resolutions to induce site-specific DSBs in living cells. This study sets the blueprint for further systematic studies of the DDR that combine vfCRISPR with time-resolved biochemical, sequencing, and imaging readouts. The use of cgRNA with other Cas9-based systems such as nickases, base editors, and prime editors may facilitate the study of single-strand break, base excision or mismatch, and flap repair, respectively. Combining vfCRISPR with subcellular photoactivation potentially enables precise genome editing with single-allele specificity and elimination of off-target activity.

REFERENCES AND NOTES

1. G. J. Knott, J. A. Doudna, *Science* **361**, 866–869 (2018).
2. S. E. Polo, S. P. Jackson, *Genes Dev.* **25**, 409–433 (2011).
3. M. van Overbeek *et al*., *Mol. Cell* **63**, 633–646 (2016).
4. B. R. Lemos *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **115**, E2040–E2047 (2018).
5. S. Lin, B. T. Staahl, R. K. Alla, J. A. Doudna, *eLife* **3**, e04766 (2014).
6. E. M. Kallimasioti-Pazi *et al*., *PLOS Biol.* **16**, e2005595 (2018).
7. L. E. Dow *et al*., *Nat. Biotechnol.* **33**, 390–394 (2015).
8. B. Zetsche, S. E. Volz, F. Zhang, *Nat. Biotechnol.* **33**, 139–142 (2015).
9. Y. Nihongaki, F. Kawano, T. Nakajima, M. Sato, *Nat. Biotechnol.* **33**, 755–760 (2015).
10. J. Hemphill, E. K. Borchardt, K. Brown, A. Asokan, A. Deiters, *J. Am. Chem. Soc.* **137**, 5642–5645 (2015).
11. P. K. Jain *et al*., *Angew. Chem. Int. Ed.* **55**, 12440–12444 (2016).
12. J. E. Dahlman *et al*., *Nat. Biotechnol.* **33**, 1159–1161 (2015).
13. D. Singh, S. H. Sternberg, J. Fei, J. A. Doudna, T. Ha, *Nat. Commun.* **7**, 12778 (2016).
14. D. Singh *et al*., *Nat. Struct. Mol. Biol.* **25**, 347–354 (2018).
15. S. H. Sternberg, B. LaFrance, M. Kaplan, J. A. Doudna, *Nature* **527**, 110–113 (2015).
16. H. Lusic, D. D. Young, M. O. Lively, A. Deiters, *Org. Lett.* **9**, 1903–1906 (2007).
17. J. C. Rose *et al*., *Nat. Methods* **14**, 891–896 (2017).
18. B. Wienert *et al*., *Science* **364**, 286–289 (2019).
19. S. Q. Tsai *et al*., *Nat. Biotechnol.* **33**, 187–197 (2015).
20. H. Yin *et al*., *Nat. Chem. Biol.* **14**, 311–316 (2018).
21. E. K. Brinkman *et al*., *Mol. Cell* **70**, 801–813.e6 (2018).
22. R. Aleksandrov *et al*., *Mol. Cell* **69**, 1046–1061.e5 (2018).
23. T. Clouaire *et al*., *Mol. Cell* **72**, 250–262.e6 (2018).
24. B. Chen *et al*., *Cell* **155**, 1479–1491 (2013).
25. H. Ma *et al*., *Nat. Biotechnol.* **34**, 528–530 (2016).
26. H. Wang *et al*., *Science* **365**, 1301–1305 (2019).
27. N. Dimitrova, Y. C. M. Chen, D. L. Spector, T. de Lange, *Nature* **456**, 524–528 (2008).
28. D. Paquet *et al*., *Nature* **533**, 125–129 (2016).

ACKNOWLEDGMENTS

We thank D. Singh for demonstrating initial feasibility on synthetic DNA, B. G. Kang and T. Dawson for selection of the monoclonal U-2 OS cell lines, T. Cottle for helping with in vitro cleavage assays, G. Seydoux for access to the Lonza 4D-Nucleofector, and W. Timp for access to the MiSeq. **Funding:** This work was supported by grants from the National Institutes of Health (R35 GM 122569 to T.H.), the National Science Foundation (PHY 1430124, EFMA 1933303 to T.H. and 1817447 to B.W.), and the Pew Charitable Trust (00030601 to B.W.). R.S.Z. was supported by an NIH Medical Scientist Training Program Award (T32 GM 007309). Y.N. was supported by Japan Society for the Promotion of Science Postdoctoral Fellowships for Research Abroad. This work was also partially supported by NIH GM123130 to Takanari Inoue. T.H. is an investigator of the Howard Hughes Medical Institute. **Author contributions:** Y.L. and T.H. conceived the project. Y.N. and Y.L. demonstrated initial feasibility in mammalian cells. R.S.Z. and Y.L. performed in vitro cleavage assays. R.S.Z., Y.L., and Y.N. performed DSB/indel kinetics measurements. R.S.Z. performed ChIP-seq and GUIDE-seq. Y.L. and S.H. performed immunofluorescence and live-cell imaging. Y.L. and Y.N. developed the monoclonal U-2 OS cell lines. X.L. performed immunoblotting. Y.L., S.R., and S.H. purified SpCas9 proteins. B.W., R.S.Z., and Y.L. developed the theoretical models. Y.L., R.S.Z., B.W., and T.H. wrote the manuscript. T.H. and B.W. supervised the project. **Competing interests:** The authors and Johns Hopkins University have filed a provisional patent application on the method of spatiotemporal control of Cas9 activities through cgRNA. **Data and materials availability:** All data associated with this study are present in the main text or the supplementary materials. All ChIP-seq data have been uploaded to the Sequence Read Archive under BioProject accession PRJNA609749, with analysis code available on GitHub (https://github.com/rogerzou/chipseq_cgRNA). mCherry-BP1-2 pLPC-Pro was obtained from Addgene (#19835) through a material transfer agreement (#39474).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1265/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S26
Tables S1 to S10
References (*29*–*38*)
Movies S1 to S3
MDAR Reproducibility Checklist

21 July 2019; resubmitted 29 December 2019
Accepted 14 April 2020
10.1126/science.aay8204

SEXUAL DIMORPHISM

A genetic mechanism for sexual dichromatism in birds

Małgorzata A. Gazda[1,2]*, Pedro M. Araújo[1,3]*, Ricardo J. Lopes[1]*, Matthew B. Toomey[4,5], Pedro Andrade[1,2], Sandra Afonso[1], Cristiana Marques[1,2], Luís Nunes[2], Paulo Pereira[1,2], Sandra Trigo[1], Geoffrey E. Hill[6], Joseph C. Corbo[4], Miguel Carneiro[1,2]†

Sexual dichromatism, a difference in coloration between males and females, may be due to sexual selection for ornamentation and mate choice. Here, we show that carotenoid-based dichromatism in *mosaic* canaries, a hybrid phenotype that arises in offspring of the sexually dichromatic red siskin and monochromatic canaries, is controlled by the gene that encodes the carotenoid-cleaving enzyme β-carotene oxygenase 2 (*BCO2*). Dichromatism in *mosaic* canaries is explained by differential carotenoid degradation in the integument, rather than sex-specific variation in physiological functions such as pigment uptake or transport. Transcriptome analyses suggest that carotenoid degradation in the integument might be a common mechanism contributing to sexual dichromatism across finches. These results suggest that differences in ornamental coloration between sexes can evolve through simple molecular mechanisms controlled by genes of major effect.

In sexually dichromatic species, males and females can differ in color or pattern (*1*, *2*). These differences in coloration arise through distinct selective pressures on the two sexes (*3*, *4*), and patterns of dichromatism can shift rapidly in response to changing social environments or predation (*5*, *6*). Whether they arise in response to selection for status signals, honest signals of quality, or simply aesthetic beauty, the genetic and molecular mechanisms that control the differences in male and female coloration remain largely unknown.

Dichromatism in birds can involve pigmentary or structural mechanisms that give rise to feather coloration. However, differences in the coloration of males and females most frequently involve red or yellow carotenoid coloration (*2*, *6*, *7*), which is also the form of ornamental coloration used in sexual signaling that is best understood (*8*, *9*). Carotenoid coloration plays a central role in mate choice and can signal social dominance (*10*). Thus, elucidating the genetic basis of carotenoid-based sexual dichromatism is important to comprehensively understand the evolution of sexual dichromatism and the selective forces that shape ornamentation in animals (*11*, *12*).

In birds, heritable differences in dichromatism are largely fixed between species (*13*). The lack of genetic differences for this trait segregating within populations complicates attempts to link genotype and phenotype. We thus took advantage of the *mosaic* breed of domesticated canaries created by an interspecific cross between the sexually dimorphic red siskin (*Spinus cucullatus*) and common canaries lacking dichromatism (*Serinus canaria*). *Mosaic* canaries are strongly dichromatic, with males accumulating more carotenoid pigment in their feathers than females (Fig. 1 and fig. S1). The *mosaic* phenotype segregates in a Mendelian fashion.

To elucidate the genetic basis of dichromatism, we conducted whole-genome sequencing of two *mosaic* breeds and compared them with four domestic breeds and one wild population of common canary (*14*). After the hybridization of red siskin with canaries, siskin alleles controlling sexual dichromatism were selected through generations of backcrossing to common canaries (Fig. 1). We therefore predict that genome sequences of *mosaic* canaries should be very similar to those of common canaries except in the region mediating dichromatism, which should be derived from the red siskin genome.

We carried out genetic differentiation [fixation index (F_{ST})], association [Cochran-Mantel-Hanzel test (CMH)], and introgression analyses [the fraction of the genome shared through introgression ($f\hat{}_d$) and the relative node depth (RND)] (*14*). These analyses revealed a clear outlier region on scaffold NW_007931177 (Fig. 2, A to D, and figs. S2 and S3), which is homologous to zebra finch chromosome 24. A second weaker signal overlapped *CYP2J19*, a gene associated with red feather coloration in canaries (*15*). Genetic differentiation is inflated at the *CYP2J19* locus by our use of breeds exhibiting both yellow and red coloration, and the signal disappears when the comparison is restricted to breeds exhibiting the same background color (fig. S4). Thus, we do not believe that *CYP2J19* plays a role in dichromatism. By contrast, differentiation, association, and introgression statistics indicate that the outlier locus on scaffold NW_007931177 is a strong candidate for controlling dichromatism in *mosaic* canaries.

Next, we increased mapping resolution at the locus on scaffold NW_007931177 by genotyping 52 variants fixed for alternative alleles between wild canaries and red siskin (Fig. 2E). Because the *mosaic* phenotype follows a recessive inheritance pattern, the expectation is that *mosaic* birds should be homozygous for a haplotype derived from the siskin genome. Consistent with this expectation, we found 12 consecutive variants homozygous for the red siskin allele in all *mosaic* canaries, defining a stretch of ~36 kb (NW_007931177:821,814 to 857,981 base pairs). This interval contained three genes: *PTS* (6-pyruvoyltetrahydropterin synthase), *BCO2* (β-carotene oxygenase 2), and *TEX12* (testis-expressed protein 12).

Given that red siskins and canaries belong to different genera (*16*), it is possible that genomic rearrangements may have occurred between the two species. We thus sequenced a red siskin individual at 6.5× coverage using

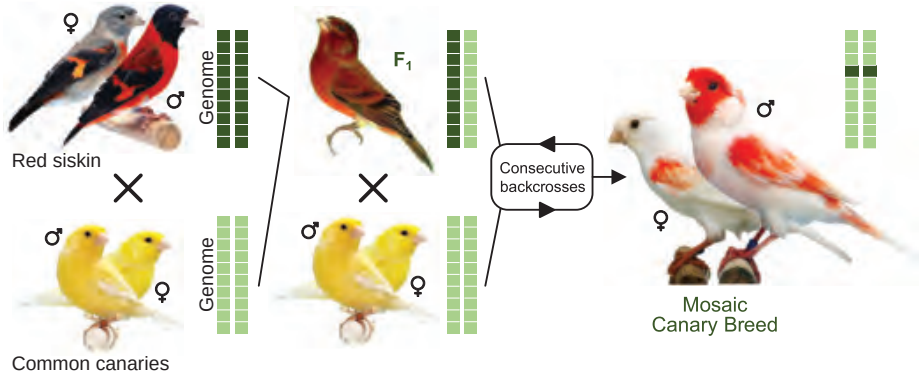


Fig. 1. Mosaic canaries were obtained through an interspecific cross. Diagram of the crosses used by breeders to obtain sexually dimorphic *mosaic* canaries from common canaries and red siskins.

[1]CIBIO/InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, Universidade do Porto, 4485-661 Vairão, Portugal. [2]Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal. [3]MARE – Marine and Environmental Sciences Centre, Department of Life Sciences, University of Coimbra, Coimbra, Portugal. [4]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA. [5]Department of Biological Science, University of Tulsa, Tulsa, OK 74104, USA. [6]Department of Biological Sciences, Auburn University, Auburn, AL 36849, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: miguel.carneiro@cibio.up.pt

long-read technology. The synteny in both species was well preserved, and no evidence was observed that genes present in the red siskin dichromatism-associated haplotype were absent from the homologous region in the common canary genome, or vice versa (fig. S5).

Sexually dimorphic phenotypes arise from differences in gene regulation between sexes (*17*, *18*). Accordingly, we measured the expression of *PTS*, *BCO2*, and *TEX12* in regenerating feather follicles from the sexually dichromatic uropygium region of *mosaic* canaries in male and female birds by quantitative polymerase chain reaction (qPCR) (Fig. 3A). We found no significant differences in expression between males and females for *PTS* or *TEX12* [Mann-Whitney *U* rank sum test (MWU), $P > 0.21$]. By contrast, we observed significantly increased expression of *BCO2* in females compared with males (MWU, $P = 0.02$).

BCO2 is a carotene-cleaving enzyme that localizes to mitochondria and catalyzes the 9′,10′ oxidative cleavage of carotenoids, an essential step in carotenoid degradation (*19*, *20*). Missense or knockout mutations in *BCO2* result in increased accumulation of carotenoids in tissues (*19*, *21*). Thus, *BCO2* represents a candidate for mediating the *mosaic* phenotype. The increased expression of *BCO2* in *mosaic* females is predicted to result in enhanced carotenoid degradation and consequent depigmentation of the integument. When we measured *BCO2* expression in the liver, another organ that plays a role in carotenoid metabolism in birds (*22*), expression levels were indistinguishable between males and females (MWU, $P > 0.80$) (Fig. 3A). This indicates that variation at the *mosaic* locus likely alters the expression of *BCO2* between sexes in a tissue-specific manner.

To further characterize *BCO2* expression patterns, we analyzed developing feather follicles of male and female birds by in situ hybridization (Fig. 3B and fig. S6). We observed *BCO2* expression in the barb ridges and barbule cells of developing white feather follicles in both sexes. We did not observe *BCO2* expression in barb ridges of carotenoid-pigmented follicles. This suggests that *BCO2* is selectively expressed in developing white feather follicles and produces the *mosaic* phenotype through the local degradation of carotenoids.

To measure allele-specific expression in regenerating feather follicles, we crossed *mosaic* canaries to common canaries and generated birds heterozygous for the dichromatism-associated siskin allele and for the common canary allele at the *BCO2* locus (Fig. 3C and table S1). In heterozygous birds, the two *BCO2* alleles are influenced by the same trans-acting regulatory elements and other environmental factors; thus, differences in their relative

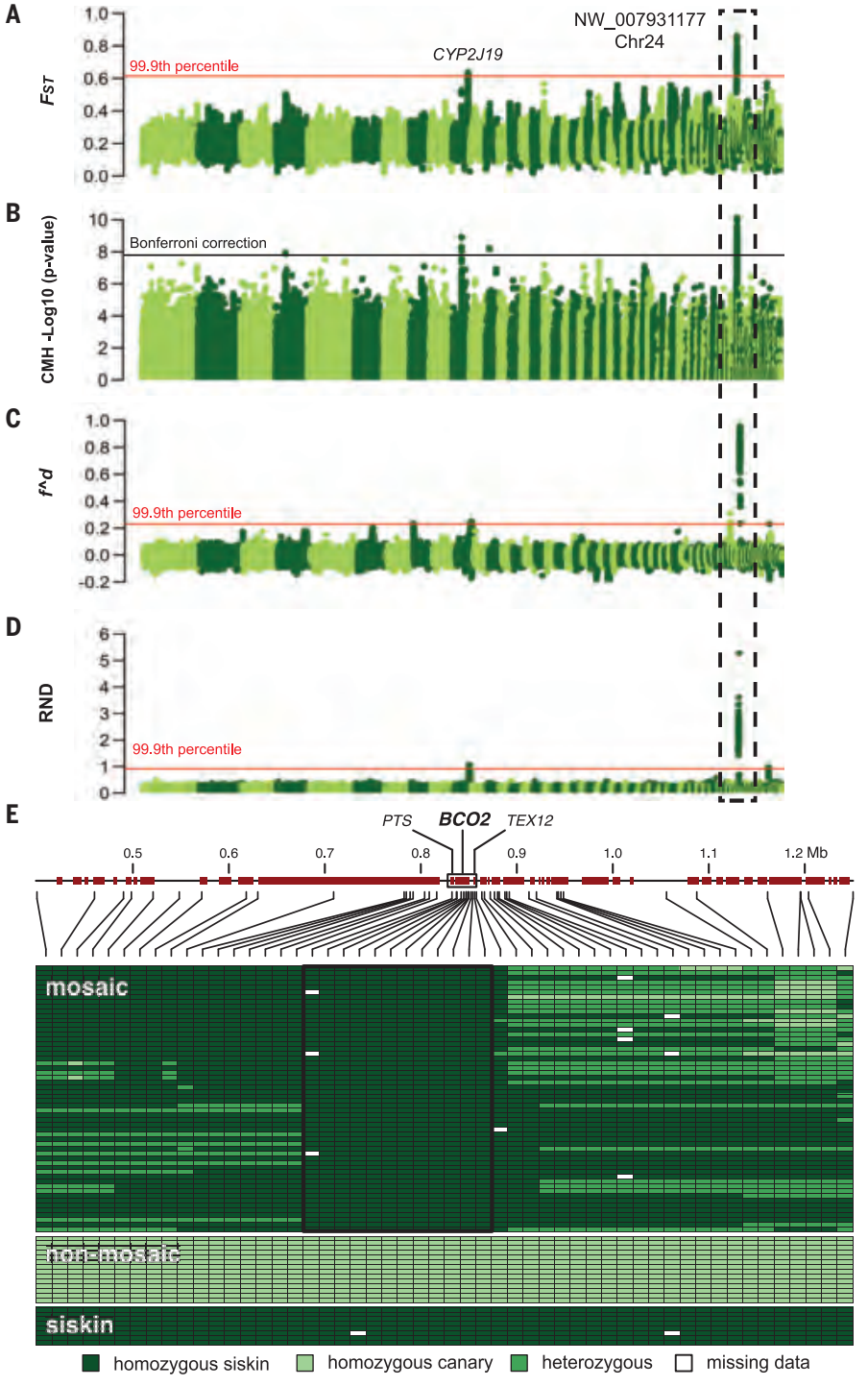


Fig. 2. Genetic mapping using whole-genome sequencing. (**A**) Average F_{ST} values between *mosaic* and non-*mosaic* canaries across the genome (20-kb windows with 5-kb steps) (**B**) −log10 values per variant of the CMH statistic configured to detect consistent differences in allele frequency between *mosaic* and non-*mosaic* canaries. (**C**) The fraction of introgression ($f\hat{}_d$) from red siskin to *mosaic* canaries summarized in nonoverlapping windows of 100 single-nucleotide polymorphisms (SNPs). (**D**) Divergence between *mosaic* and non-*mosaic* canaries summarized across the genome using RND. Dots represent RND values in nonoverlapping windows of 10,000 polymorphic and nonpolymorphic positions passing filters. In (A) to (D), the 99.9th percentile of the empirical distribution and the significance threshold after Bonferroni correction are shown by red and black horizontal lines, respectively. (**E**) Genotyping across the candidate region on scaffold NW_007931177. Each column represents one SNP, and each row represents one individual. Shades of green indicate positions homozygous for the siskin allele, homozygous for the canary allele, and heterozygous. White indicates missing data. Protein-coding genes are indicated by red boxes.

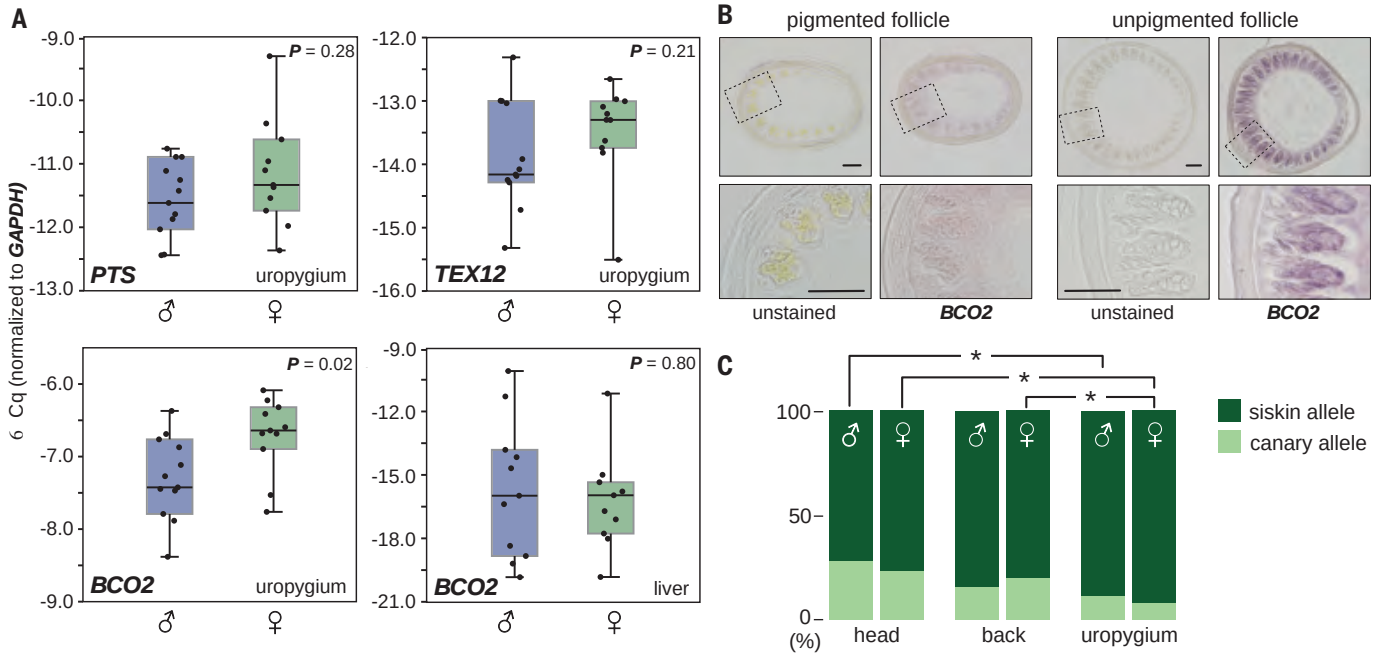


Fig. 3. Gene expression analysis in *mosaic* canaries. (**A**) qPCR measurements of *BCO2*, *PTS*, and *TEX12* normalized to *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase) in the uropygium skin and liver of *mosaic* canaries. Relative expression (ΔCq) was obtained by averaging quantification cycle (Cq) values of three technical replicates. Boxes represent the 25 and 75 percent quartiles, horizontal lines inside the box mark the median, and the short horizontal lines ("whiskers") indicate the minimal and maximal values. (**B**) Unstained sections and in situ hybridization of *BCO2* in regenerating feather follicles that express or lack carotenoid pigments in the developing barb ridges. The bottom row shows magnified views of the outlined areas in the images in the top row. Scale bars, 50 μm. (**C**) Relative expression (%) of the red siskin and canary alleles in regenerating feather follicles of heterozygous individuals. Values are averages of two amplicons (table S1). Significant comparisons using FET are denoted with an asterisk.

expression should be due to cis-acting regulatory elements (i.e., an enhancer or promoter), which affect gene expression in an allele-specific manner (*23*). We found preferential expression of the siskin allele over the canary allele in both sexes and in the three feather tracts sampled [Fisher's exact test (FET), $P < 10^{-16}$]. This difference is likely due to a more active cis-regulatory element on the siskin haplotype and likely explains why the integument of *mosaic* canaries in both sexes exhibits less carotenoid pigmentation compared with common canaries (Fig. 1).

Our sample size of one male and female is small, and the relative expression of both alleles was not evidently different between sexes (FET, $P > 0.05$). However, we did observe significant differences among the three feather tracts (FET, $P < 0.05$) (Fig. 3C), suggesting that trans-acting regulators in the canary genomic background might regionally modulate *BCO2* expression across the integument. Trans-acting regulators of *BCO2* expression and/or additional genes located elsewhere in the genome, which can alter the rate at which carotenoids are deposited or degraded, could explain why *mosaic* canaries and red siskins exhibit different carotenoid pigmentation patterns (Fig. 1) despite sharing identical DNA sequences at the *BCO2* locus (Fig. 2E). Our results show that dichromatic phenotypes can be produced by genes of large effect; however, they also suggest that additional genetic modifiers might be involved in the fine-tuning of the dichromatism observed in nature.

The canary and red siskin BCO2 proteins also differ at two amino acid positions. However, these substitutions are found in bird species lacking sexual dichromatism or carotenoid pigmentation (fig. S7). Thus, a functional role for these differences is unlikely. Overall, our expression studies suggest that sexual dichromatism in *mosaic* canaries arises because of differences in the activity of *BCO2* throughout the integument. The lack of ornamental coloration in female *mosaic* canaries seems to be estrogen-dependent, because reproductively senescent and ovariectomized females develop a color pattern similar to that of males (*24*). These observations suggest the presence of siskin-derived, hormone-responsive regulatory elements within the introgressed haplotype. These putative regulatory elements have yet to be identified.

To test whether the mechanisms for sexual dichromatism that we uncovered in *mosaic* canaries are present in wild bird species, we examined gene expression in the developing feathers of three species of finches that vary in the extent of carotenoid-based sexual dichromatism (Fig. 4A): common canaries (*S. canaria*) exhibiting wild-type coloration, which exhibit slight sexual dichromatism; the European serin (*Serinus serinus*), the sister species of canaries, which displays more pronounced sexual dichromatism; and the house finch (*Haemorhous mexicanus*), a species in which males display bright red or yellow colors but females are nearly devoid of colorful carotenoids in their plumage. We sampled the same regions of the integument in all three species (chest and belly) and profiled gene expression by RNA sequencing (*14*).

We tested if the degree of sexual dichromatism was correlated with gene expression divergence between males and females. We found that feather patches that differed more strongly in carotenoid pigmentation between sexes (chest and belly in serin and chest in house finch) had a larger number of differentially expressed genes (DEGs) (Fig. 4B), demonstrating that sexual dichromatism correlates with increased sex-biased gene expression.

We also compared gene expression between sexes, sister species, or patches within species showing pronounced differences in levels of carotenoid pigmentation (Fig. 4C). Because the transcriptomes of males and females differ in most tissues (*17*), many of the expression differences observed in our dataset should have no causal relationship with pigmentation differences. We thus reasoned that DEGs shared among the three types of contrasts would be promising candidates mediating dichromatism. Of the DEGs from our total dataset, only 12 genes met this criterion (Fig. 4D and table S2), including *BCO2*.

A closer examination revealed that several aspects of *BCO2* expression varied predictably

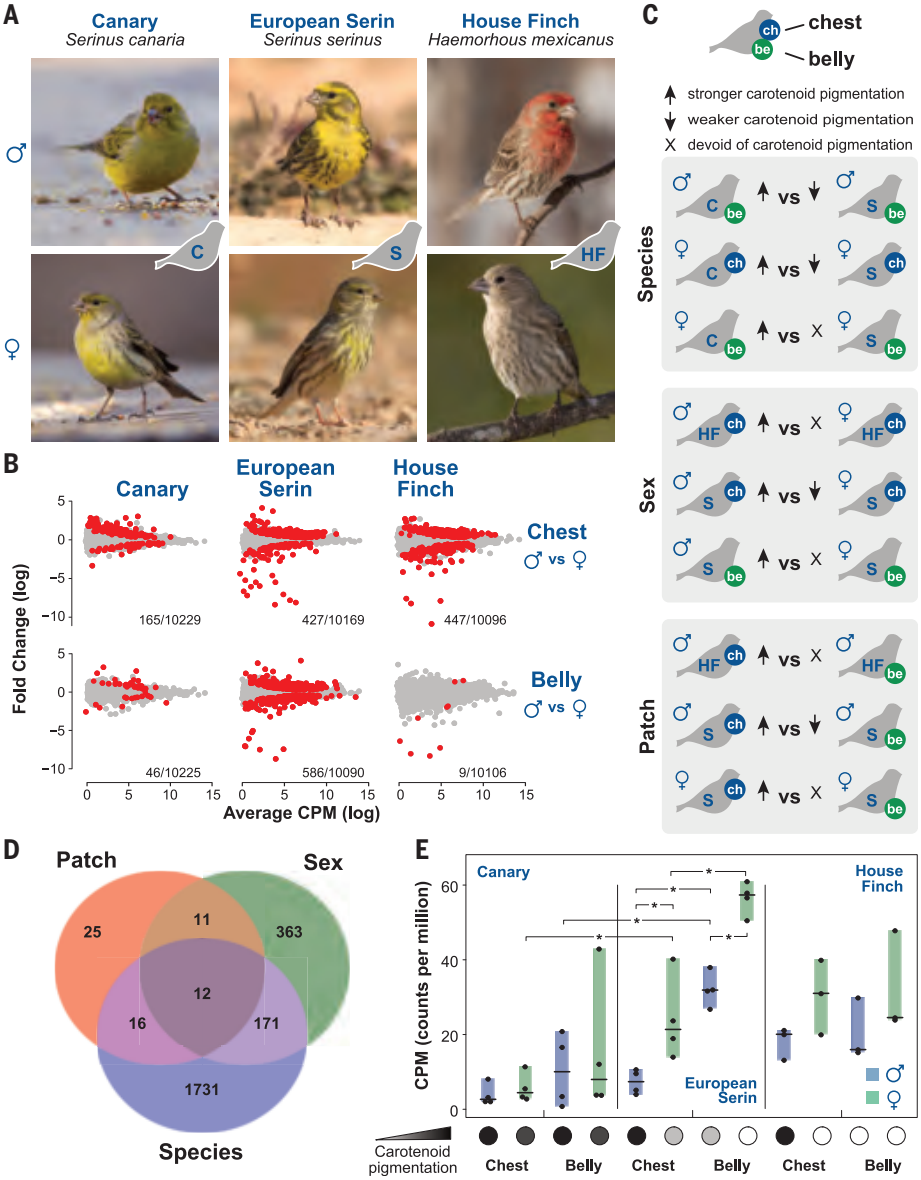


Fig. 4. Transcriptomics along a continuum of sexual dichromatism. (**A**) Representative pictures of male and female canaries (C), European serin (S), and house finch (HF). (**B**) Scatterplots of log-fold change (*y* axis) and log-CPM (counts per million) (*x* axis) in comparisons between males and females. Significant genes are depicted as red dots. For each comparison, the number of DEGs versus the total number of genes in the transcriptome is given at the bottom of the graph. (**C**) Nine different comparisons characterized by pronounced differences in the intensity of carotenoids, which include contrasts between sexes, species, and patches [belly (be) and chest (ch)]. (**D**) DEGs among the types of contrasts defined in (B). (**E**) Patterns of *BCO2* expression in the three finch species. Comparisons where *BCO2* was found significantly differentially expressed by the three methods implemented are denoted with an asterisk (*14*). Carotenoid intensity is indicated by gray scale.

with plumage carotenoid content in comparisons involving the European serin (Fig. 4E). First, serin females exhibit less carotenoid pigmentation than serin males and expressed *BCO2* at higher levels. Second, *BCO2* expression was higher in serin females compared with canary females, whereas the latter exhibit more marked carotenoid pigmentation. Finally, both male and female serins show lower expression of *BCO2* in feather patches exhibiting stronger carotenoid pigmentation. These findings suggest that *BCO2* plays a role in dichromatism in the European serin. *BCO2* expression was largely uncorrelated with the levels of carotenoid pigmentation in comparisons involving the house finch (Fig. 4E), which suggests that finches may use alternative molecular mechanisms to produce sexual dichromatism.

Genetic studies of sexual dichromatism may reveal the molecular mechanisms that enable the expression of differential male and female traits from a single shared genome. Here, we show that differences in carotenoid pigmentation between sexes of *mosaic* canaries are controlled by a single genomic region of red siskin origin. This region contains *BCO2*, a candidate gene for sex-specific pigmentation in birds displaying sexually divergent carotenoid-based coloration. The simplicity of this genetic mechanism may help explain the evolutionary lability of sexual dichromatism and why carotenoid pigmentation is the coloration mechanism most commonly associated with sexual dichromatism in birds. Sexual dichromatism in carotenoid coloration has been proposed to result from sex-specific differences in ingestion, absorption, metabolism, or transportation of carotenoids (*25*). Our observations, however, suggest that selective degradation of carotenoids at different rates in peripheral tissues may be an important mechanism for differences in carotenoid coloration in males and females. These findings add to a growing body of evidence that the *BCO2* locus is a genomic hotspot for the evolution of carotenoid-based pigmentation across multiple tissues and vertebrates (*26*–*28*), which our study extends to sexual dichromatism.

REFERENCES AND NOTES

1. T. Amundsen, *Trends Ecol. Evol.* **15**, 149–155 (2000).
2. A. V. Badyaev, G. E. Hill, *Annu. Rev. Ecol. Evol. Syst.* **34**, 27–49 (2003).
3. J. Dale, C. J. Dey, K. Delhey, B. Kempenaers, M. Valcu, *Nature* **527**, 367–370 (2015).
4. A. E. Johnson, J. Jordan Price, S. Pruett-Jones, *Ecol. Evol.* **3**, 3030–3046 (2013).
5. T. Price, G. L. Birch, *Auk* **113**, 842–848 (1996).
6. A. V. Badyaev, G. E. Hill, *Biol. J. Linn. Soc. London* **69**, 153–172 (2000).
7. D. A. Gray, *Am. Nat.* **148**, 453–480 (1996).
8. R. J. Weaver, E. S. A. Santos, A. M. Tucker, A. E. Wilson, G. E. Hill, *Nat. Commun.* **9**, 73 (2018).
9. R. J. Weaver, R. E. Koch, G. E. Hill, *Philos. Trans. R. Soc. London B Biol. Sci.* **372**, 20160343 (2017).
10. P. A. Svensson, B. B. M. Wong, *Behaviour* **148**, 131–189 (2011).
11. J. A. Coyne, E. H. Kay, S. Pruett-Jones, *Evolution* **62**, 214–219 (2008).
12. E. R. Funk, S. A. Taylor, *Auk* **136**, ukz048 (2019).
13. K. J. Burns, *Evolution* **52**, 1219–1224 (1998).
14. Materials and methods are available as supplementary materials.
15. R. J. Lopes *et al.*, *Curr. Biol.* **26**, 1427–1434 (2016).
16. D. Zuccon, R. Prŷs-Jones, P. C. Rasmussen, P. G. P. Ericson, *Mol. Phylogenet. Evol.* **62**, 581–596 (2012).
17. H. Ellegren, J. Parsch, *Nat. Rev. Genet.* **8**, 689–698 (2007).
18. T. M. Williams, S. B. Carroll, *Nat. Rev. Genet.* **10**, 797–804 (2009).
19. J. Amengual *et al.*, *FASEB J.* **25**, 948–959 (2011).
20. C. Dela Seña *et al.*, *J. Biol. Chem.* **291**, 14609–14619 (2016).
21. M. A. Gazda *et al.*, *Mol. Biol. Evol.* **37**, 1317–1328 (2020).
22. E. del Val *et al.*, *Naturwissenschaften* **96**, 797–801 (2009).
23. P. J. Wittkopp, B. K. Haerum, A. G. Clark, *Nature* **430**, 85–88 (2004).
24. O. Perez-Beato, *Fundamentals of Color Genetics in Canaries: Reproduction and Control* (RoseDog Books, Pittsburgh, 2008).
25. K. J. McGraw, G. E. Hill, R. Stradi, R. S. Parker, *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **131**, 261–269 (2002).
26. J. Eriksson *et al.*, *PLOS Genet.* **4**, e1000010 (2008).
27. D. P. L. Toews *et al.*, *Curr. Biol.* **26**, 2313–2318 (2016).
28. P. Andrade *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5633–5642 (2019).

29. M. Carneiro *et al*., A genetic mechanism for sexual dichromatism in birds. Dryad (2020); doi:10.5061/dryad.80gb5mknh.
30. P. Pereira, PJADPereira/Dichromatism: Dichromatism. Zenodo (2020); doi:10.5281/zenodo.3648365.

ACKNOWLEDGMENTS

We thank A. C. Lemo (mosaic and siskin), C. Loukeris (canaries from Fig. 1 and F_1 hybrid), S. Cherrug (canaries from Fig. 4), M. Carlo (male house finch), B. Richmond (serin male), and D. Williams (serin female) for providing photographs. We also thank Ibanidis Lda (Versele-Laga distributor) for food supplements and cages. **Funding:** This work was supported by the Fundação para a Ciência e Tecnologia (FCT) through POPH-QREN funds from the European Social Fund and Portuguese MCTES (IF/00283/2014/CP1256/CT0012 and CEECINST/00014/2018); by National Funds [Transitory Norm contract to R.J.L. (DL57/2016/CP1440/CT0006)]; by research fellowships attributed to M.A.G. (PD/BD/114042/2015), P.A. (PD/BD/114028/2015), and P.P. (PD/BD/128492/2017) in the scope of the Biodiversity, Genetics, and Evolution (BIODIV) Ph.D. program; and by the project "PTDC/BIA-EVL/31569/2017 - NORTE -01-0145-FEDER-30288," cofunded by NORTE2020 through Portugal 2020 and FEDER Funds and by National Funds through FCT. M.B.T. received support from the University of Tulsa. **Author contributions:** M.C., J.C.C., and G.E.H. conceived the study with input from M.A.G., P.M.A., R.J.L., and M.B.T.; M.C. coordinated the study; M.A.G., P.M.A., R.J.L., M.B.T., P.A., S.A., C.M., L.N., P.P., and S.T. performed research; M.A.G., P.M.A., R.J.L., P.A., L.N., P.P., and M.C. analyzed data; M.C., J.C.C., G.E.H., M.A.G., and M.B.T. wrote the paper with input from all other authors. All authors approved the manuscript before submission. **Competing interests:** None declared. **Data and materials availability:** Whole-genome sequencing, RNA sequencing, allelic imbalance, and nanopore sequencing data are available in the Sequence Read Archive (www.ncbi.nlm.nih.gov/sra) under BioProject PRJNA591356. SNP calling, genotyping data, de novo transcriptome assemblies, and full lists of differentially expressed genes are available from the Dryad digital repository (*29*). Computer code is available at Zenodo (*30*).

SUPPLEMENTARY MATERIAL

science.sciencemag.org/content/368/6496/1270/suppl/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S6
References (*31*–*61*)
MDAR Reproducibility Checklist

2 November 2019; accepted 10 April 2020
10.1126/science.aba0803

CORONAVIRUS

A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2

Yan Wu[1,2]*†, Feiran Wang[3,4]*, Chenguang Shen[3,5]*, Weiyu Peng[3,6]*, Delin Li[3,5,7]*, Cheng Zhao[3,8], Zhaohui Li[3,9], Shihua Li[3], Yuhai Bi[3,10], Yang Yang[5], Yuhuan Gong[3,10], Haixia Xiao[7], Zheng Fan[3], Shuguang Tan[3], Guizhen Wu[11], Wenjie Tan[11], Xuancheng Lu[12], Changfa Fan[13], Qihui Wang[3], Yingxia Liu[5], Chen Zhang[1], Jianxun Qi[3], George Fu Gao[3]†, Feng Gao[7]†, Lei Liu[5]†

Neutralizing antibodies could potentially be used as antivirals against the coronavirus disease 2019 (COVID-19) pandemic. Here, we report isolation of four human-origin monoclonal antibodies from a convalescent patient, all of which display neutralization abilities. The antibodies B38 and H4 block binding between the spike glycoprotein receptor binding domain (RBD) of the virus and the cellular receptor angiotensin-converting enzyme 2 (ACE2). A competition assay indicated different epitopes on the RBD for these two antibodies, making them a potentially promising virus-targeting monoclonal antibody pair for avoiding immune escape in future clinical applications. Moreover, a therapeutic study in a mouse model validated that these antibodies can reduce virus titers in infected lungs. The RBD-B38 complex structure revealed that most residues on the epitope overlap with the RBD-ACE2 binding interface, explaining the blocking effect and neutralizing capacity. Our results highlight the promise of antibody-based therapeutics and provide a structural basis for rational vaccine design.

Coronavirus disease 2019 (COVID-19) caused by the novel COVID-19 virus has become a pandemic. The virus has spread worldwide, causing fever, severe respiratory illness, and pneumonia (*1*, *2*). Phylogenetic analysis indicates that the virus is closely related to severe acute respiratory syndrome coronavirus (SARS-CoV) (*3*–*5*), but it appears to be more easily transmitted from person to person than SARS-CoV (*6*). To date, no specific drugs or vaccines are available for COVID-19.

The COVID-19 virus belongs to the betacoronavirus genus, which includes five pathogens that infect humans (*7*, *8*). Among them, SARS-CoV and Middle East respiratory syndrome coronavirus (MERS-CoV) are two highly pathogenic viruses. As with other coronaviruses, the spike (S) glycoprotein homotrimer on the COVID-19 virus surface plays an essential role in receptor binding and virus entry. The S protein is a class I fusion protein—each S protomer consists of S1 and S2 domains (*9*), with the receptor binding domain (RBD) located within the S1 domain (*8*). Previous studies have revealed that the COVID-19 virus, similarly to SARS-CoV, uses the angiotensin-converting enzyme 2 (ACE2) receptor for cell entry (*3*, *10*–*13*). Numerous neutralizing antibodies have been found to target the RBDs of SARS-CoV or MERS-CoV (*14*–*16*). Therefore, screening for neutralizing antibodies that target the COVID-19 virus RBD is a priority.

We expressed COVID-19 virus RBD protein as bait to isolate specific single memory B cells from COVID-19 patient peripheral blood mononuclear cells (PBMCs). The variable regions encoding the heavy and light chains were each amplified from separate single B cells and then cloned into a pCAGGS vector with the constant region to produce immunoglobulin G1 (IgG1) antibodies, as described previously (*17*). Seventeen paired B cell clones were amplified, three of which were identical (B5, B59, and H1). To identify the antibody binding abilities, the plasmids containing the paired heavy and light chains were cotransfected into human embryonic kidney–293T (HEK 293T) cells for monoclonal antibody (mAb) production. The supernatants were then screened for binding to the RBD by biolayer interferometry (BLI). An irrelevant anti–severe fever with thrombocytopenia syndrome virus Gn antibody and a SARS-specific antibody were used as controls (*18*). The supernatants from four different antibodies (B5, B38, H2, and H4) bound to COVID-19 virus RBD but not to SARS-CoV RBD (fig. S1), suggesting that the epitopes of the two RBDs are immunologically distinct. The usage of heavy chain (V_H) and light chain (V_L) variable genes in these four antibodies is listed in table S1.

The dissociation constants (K_d) for the four antibodies binding to COVID-19 virus RBD, measured using surface plasmon resonance (SPR), ranged from 10^{-7} to 10^{-9} M (Fig. 1, A to D). We next studied the neutralizing activities of these four antibodies against COVID-19 virus (the BetaCoV/Shenzhen/SZTH-003/2020 strain). All four antibodies exhibited neutralizing

[1]Department of Pathogen Microbiology, School of Basic Medical Sciences, Capital Medical University, Beijing, China. [2]Research Network of Immunity and Health (RNIH), Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing, China. [3]CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences (CAS), Beijing, China. [4]School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, China. [5]Shenzhen Key Laboratory of Pathogen and Immunity, Shenzhen Third People's Hospital, Shenzhen, China. [6]College of Veterinary Medicine, China Agricultural University, Beijing, China. [7]Laboratory of Protein Engineering and Vaccines, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences (CAS), Tianjin, China. [8]Shanxi Academy of Advanced Research and Innovation, Taiyuan, China. [9]University of Chinese Academy of Sciences, Beijing, China. [10]Center for Influenza Research and Early Warning, Chinese Academy of Sciences (CASCIRE), Beijing, China. [11]NHC Key Laboratory of Biosafety, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China. [12]Laboratory Animal Center, Chinese Center for Disease Control and Prevention, Beijing, China. [13]Division of Animal Model Research, Institute for Laboratory Animal Resources, National Institutes for Food and Drug Control, Beijing, China.
*These authors contributed equally to this work.
†Corresponding author. Email: wuy@biols.ac.cn (Y.W.); gaofeng@tib.cas.cn (F.G.); gaof@im.ac.cn (G.F.G.); liulei3322@aliyun.com (L.L.)

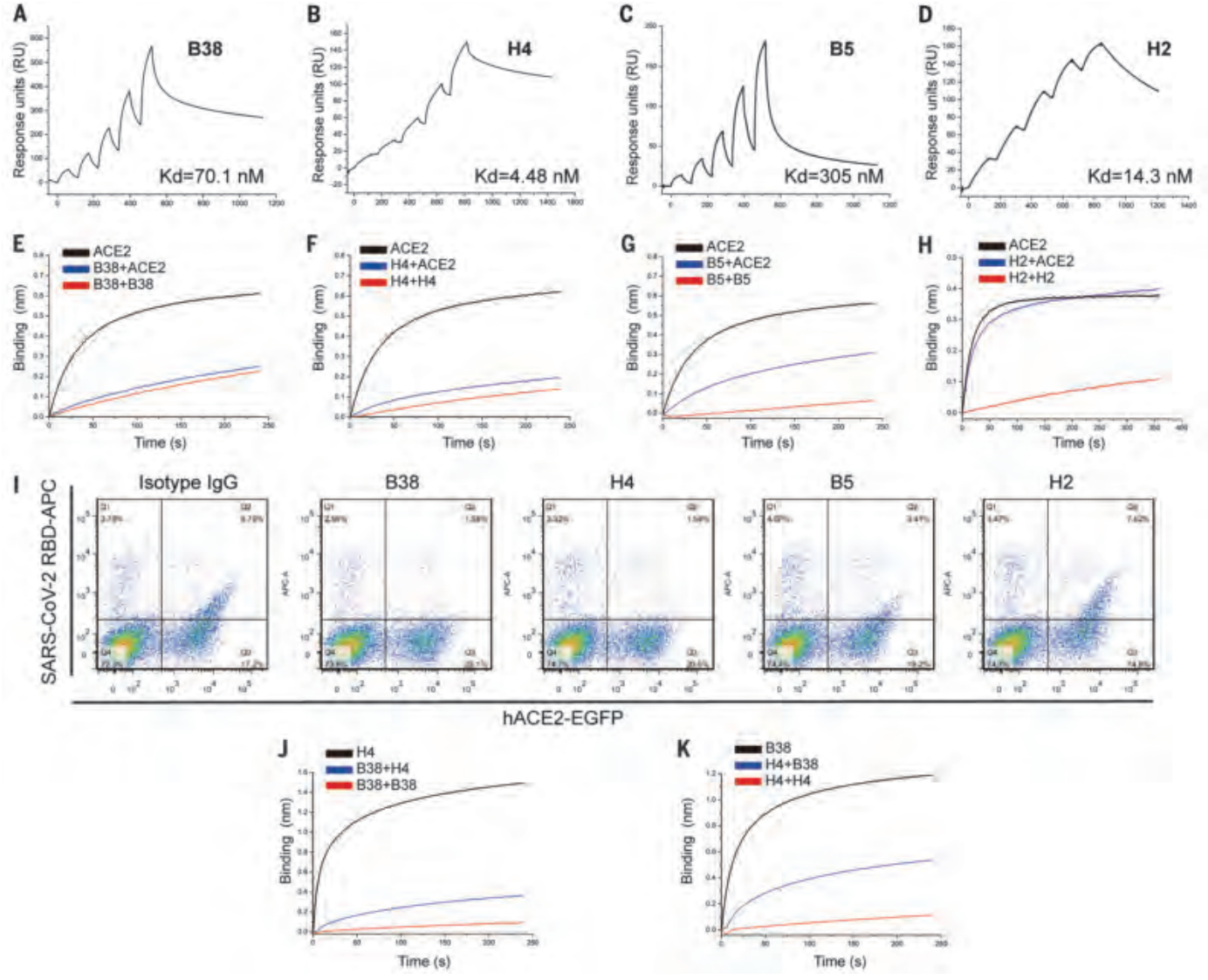


Fig. 1. Characterization of COVID-19 virus–specific neutralizing antibodies. (**A** to **D**) The binding kinetics between four antibodies (B38, H4, B5, and H2) and COVID-19 virus RBD were measured using a single-cycle Biacore 8K system. (**E** to **H**) Competition binding to the COVID-19 virus RBD between antibodies and ACE2 was measured by BLI. Immobilized biotinylated COVID-19 virus RBD (10 μg/ml) was saturated with antibodies and then flowed with corresponding antibody in the presence of 300 nM soluble ACE2 (blue) or without ACE2 (red). As a control, the immobilized biotinylated RBD was flowed with buffer and then flowed with the equal molar concentration of ACE2 (black). The graphs show binding patterns after antibody saturation. (**I**) hACE2–enhanced green fluorescent protein (EGFP) was expressed on the HEK293T cell surface, and the cells were stained with 200 ng/ml COVID-19 virus RBD his-tag proteins preincubated with isotype IgG, B38, H4, B5, or H2. The percentages of anti-his-tag APC^+ (allophycocyanin) cells and $EGFP^+$ cells were calculated. (**J** and **K**) Competition binding to COVID-19 virus RBD between B38 and H4 was measured by BLI. Immobilized COVID-19 virus RBD (10 μg/ml) was saturated with 300 nM of the first antibody and then flowed with equal molar concentration of the first antibody in the presence of (blue) or without (red) the second antibody. Equal molar concentration of the second antibody was flowed on the immobilized RBD as a control (black). The graphs show binding patterns after saturation of the first antibody.

activities, with median inhibitory concentration (IC_{50}) values ranging from 0.177 to 1.375 μg/ml (Fig. 2, A to D). A cocktail of B38 and H4 exhibited synergetic neutralizing ability, even in the presence of a higher virus titer (Fig. 2E).

To evaluate the ability of each antibody to inhibit binding between RBD and ACE2, we performed a competition assay using BLI and a blocking assay using fluorescence-activated cell sorting (FACS). For the BLI assay, streptavidin biosensors labeled with biotinylated RBD were saturated with antibodies, and then the test antibodies were flowed through in the presence of soluble ACE2. B38 and H4 showed complete competition with ACE2 for binding to RBD. In contrast, B5 displayed partial competition, whereas H2 did not compete with ACE2 for RBD binding (Fig. 1, E to H). The blocking assay by FACS presented the same result (Fig. 1I). To determine whether B38 and H4 target the same epitope, we performed an epitope competition assay by BLI. The nickel–nitrilotriacetic acid sensor labeled with the RBD was saturated with B38 IgG, and H4 IgG was flowed through, or the reverse (sensor saturated with H4 IgG, and B38 IgG flowed through). Although RBD was saturated with the first antibody, the second antibody could still bind to RBD, but with some inhibition. This suggests that B38 and H4 recognize different epitopes on RBD with partial overlap (Fig. 1, J and K).

To explore the protection efficacy of B38 and H4 against challenge with COVID-19 virus

in vivo, hACE2 transgenic mice were administered a single 25 mg/kg dose of B38 or H4 12 hours after viral challenge. The body weight of the B38 group decreased slowly and recovered at 3 days postinfection (dpi) compared with the phosphate-buffered saline (PBS) control group and the H4 group (Fig. 3A). The number of viral RNA copies in the lung were also measured at 3 dpi. The RNA copies of both the B38 group and the H4 group were significantly lower than those of the PBS group, with a reduction of 3.347 and 2.655 logs, respectively (Fig. 3B). These results show the same trends as the neutralization abilities. Histopathological examination indicated that severe bronchopneumonia

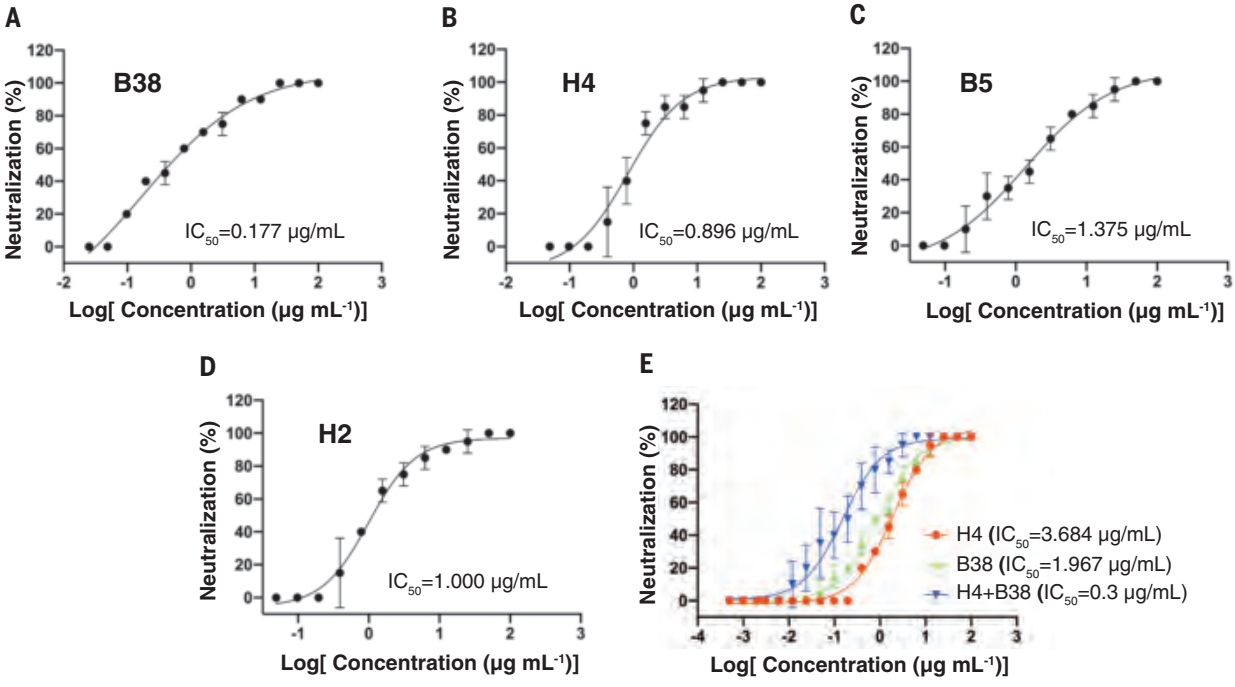
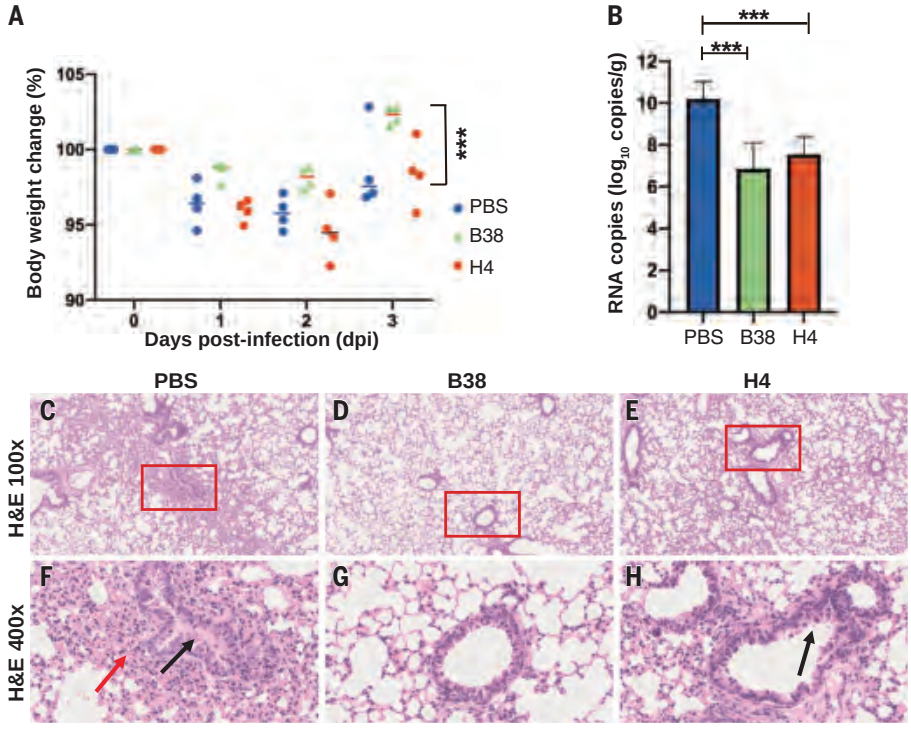


Fig. 2. Four antibodies can effectively neutralize COVID-19 virus, and two of them exhibit additive inhibition effect. The mixtures of COVID-19 virus and serially diluted antibodies were added to Vero E6 cells. After 5 days of incubation, IC_{50} values were calculated by fitting the cytopathic effect from serially diluted antibody to a sigmoidal dose-response curve. Medium containing 100 and 200 times the median tissue culture infectious dose of COVID-19 virus was used for testing the neutralizing abilities of individual antibody (**A** to **D**) and cocktail antibodies (**E**), respectively.

Fig. 3. The protection efficiency of mAbs in hACE2 mice model after infection with COVID-19 virus. (**A**) Body weight loss was recorded for PBS, B38 treatment, and H4 treatment groups (for all groups, $n = 4$ mice). All the mice were challenged intranasally with COVID-19 virus, and a 25 mg/kg dose of antibodies was injected (intraperitoneally) 12 hours after infection. Equal volume of PBS was used as a control. The weight loss was recorded over 3 days, and a significant difference could be observed between the B38 group and the PBS group (unpaired *t* test, ***$P < 0.001$). (**B**) The virus titer in lungs of three groups was determined at 3 dpi by real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR). The mAb treatment group reduced the viral load in the lungs of mice (unpaired *t* test, ***$P < 0.001$). (**C** to **H**) Representative histopathology of the lungs in COVID-19 virus–infected hACE2 mice (3 dpi). Severe bronchopneumonia and interstitial pneumonia was observed in the PBS group [(C) and (F)], with edema and bronchial epithelial cell desquamation (black arrow) and infiltration of lymphocytes within alveolar spaces (red arrow). Mild bronchopneumonia was observed in the H4 group [(E) and (H)], whereas no lesions were observed in the B38 group [(D) and (G)]. The images and areas of interest (red boxes) are magnified 100× and 400×, respectively.



and interstitial pneumonia could be observed in the mice of the PBS control group, with edema and bronchial epithelial cell desquamation and infiltration of lymphocytes within alveolar spaces (Fig. 3, C and F). Mild bronchopneumonia was observed in the H4 group (Fig. 3, E and H), whereas no lesions were observed in the B38 group (Fig. 3, D and G).

As is consistent with the binding affinity between RBD and B38 or H4, stable complexes were obtained in both RBD-B38 and RBD-H4 mixtures (fig. S2). The complex crystal structure of RBD-B38 Fab was solved at 1.9-Å resolution (table S2). Three complementarity-determining regions (CDRs) on the heavy chain and two

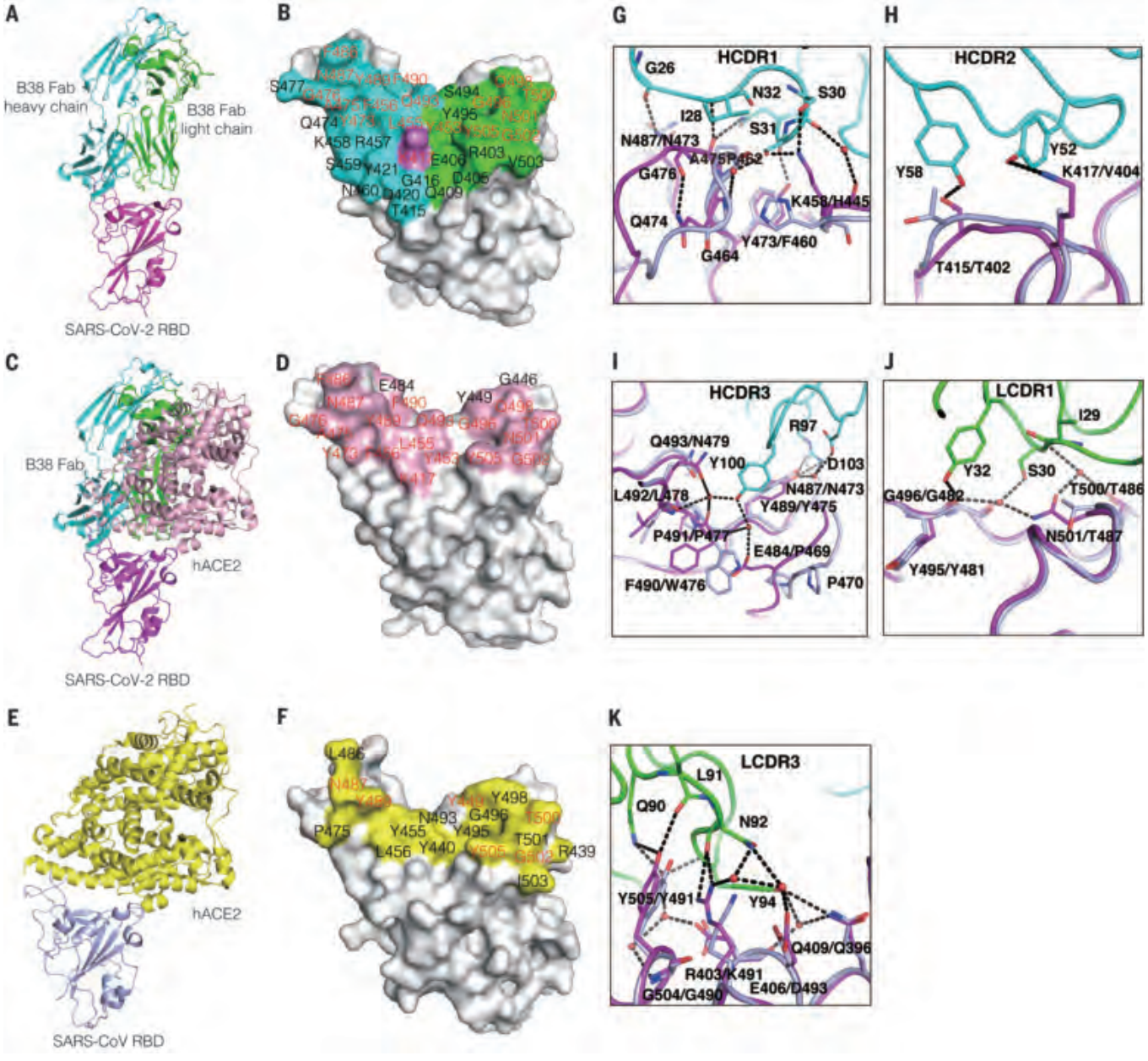


Fig. 4. Structural analysis of B38 and COVID-19 virus RBD complex and the epitope comparison between B38 and hACE2. (**A**) The overall structure of B38 Fab and COVID-19 virus RBD. The B38 heavy chain (cyan), light chain (green), and COVID-19 virus RBD (magenta) are shown in cartoon representation. (**B**) The epitope of B38 is shown in surface representation. The contact residues by heavy chain, light chain, or both are colored in cyan, green, and magenta, respectively. The residues on RBD involved in both B38 and hACE2 binding are labeled in red. (**C**) Superimposition of RBD-B38 and RBD-hACE2 [Protein Data Bank (PDB) ID 6LZG]. All molecules are shown in cartoon representation, with the same colors as in (A). hACE2 is colored in light pink. (**D**) The residues involved in hACE2-RBD binding are highlighted in light pink. The residues on RBD involved in both B38 and hACE2 binding are labeled in red. (**E**) The complex structure of SARS-CoV RBD (light blue) and hACE2 (yellow) (PDB ID 2AJF). (**F**) The residues in contact with hACE2 are colored in yellow. The residues are numbered according to SARS-CoV RBD. The residues involved in hACE2 binding of two RBDs are labeled in red. (**G** to **I**) The detailed interactions between COVID-19 virus RBD and CDR loops of the heavy chain. (**J** and **K**) The detailed interactions between COVID-19 virus RBD and CDR loops of the light chain. The residues are shown in stick representation, with the same colors as in (C). The water molecules are shown as red spheres. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; D, Asp; E, Glu; F, Phe; G, Gly; I, Ile; K, Lys; L, Leu; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

CDRs on the light chain are involved in interaction with RBD (Fig. 4, A, B, and G to K). The buried surface area of heavy and light chains on the epitope is 713.9 and 497.7 Å, respectively. There are 36 residues in the RBD involved in the interaction with B38, in which 21 residues and 15 residues interact with heavy and light chains, respectively (table S3 and Fig. 4B). Sequence alignment indicates that only 15 of the 36 residues in the epitope (defined as residues buried by B38) are conserved between COVID-19 virus and SARS-CoV (Fig. 4, D to F, and fig. S3). Notably, most contacts in the interface between B38 and RBD are hydrophilic interactions (table S4). Water molecules play an important role in the binding between COVID-19 RBD and B38 (Fig. 4, G and I to K). These differences explain the B38-specific binding to the COVID-19 virus rather than SARS-CoV.

To explore the structural basis for B38 blocking the interaction between COVID-19 virus RBD and ACE2, the complex structures of RBD–B38-Fab and RBD-hACE2 were superimposed. Both the V_H and V_L of B38 would sterically hinder ACE2 binding (Fig. 4C). Notably, the RBDs in B38-bound form and hACE2-bound form have no notable conformational differences, with a Cα root mean square deviation of 0.489 Å (for 194 atoms). Further analysis indicated that 18 of the 21 amino acids on the RBD are involved in binding both B38 and ACE2 (Fig. 4D), which explains why B38 abolishes the binding between COVID-19 virus RBD and the receptor.

As the COVID-19 outbreak continues to spread, characterization of the epitopes on the COVID-19 virus RBD will provide valuable information for vaccine development. Furthermore, the molecular features of the neutralizing antibody targeting epitopes are helpful for the development of small-molecule or peptide drugs and inhibitors. The neutralizing antibodies themselves are also promising candidates for prophylactic and therapeutic treatment against the COVID-19 virus.

REFERENCES AND NOTES

1. N. Zhu *et al*., *N. Engl. J. Med.* **382**, 727–733 (2020).
2. C. Wang, P. W. Horby, F. G. Hayden, G. F. Gao, *Lancet* **395**, 470–473 (2020).
3. P. Zhou *et al*., *Nature* **579**, 270–273 (2020).
4. W. Tan *et al*., *China CDC Weekly* **2**, 61–62 (2020).
5. R. Lu *et al*., *Lancet* **395**, 565–574 (2020).
6. J. F. Chan *et al*., *Lancet* **395**, 514–523 (2020).
7. F. Wu *et al*., *Nature* **579**, 265–269 (2020).
8. G. Lu, Q. Wang, G. F. Gao, *Trends Microbiol.* **23**, 468–478 (2015).
9. D. M. Knipe, P. M. Howley, Eds., *Fields Virology* (Lippincott Williams & Wilkins, ed. 6, 2013).
10. A. C. Walls *et al*., *Cell* **181**, 281–292.e6 (2020).
11. M. Hoffmann *et al*., *Cell* **181**, 271–280.e8 (2020).
12. Q. Wang *et al*., *Cell* **181**, 894–904.e9 (2020).
13. D. Wrapp *et al*., *Science* **367**, 1260–1263 (2020).
14. Y. Zhou, Y. Yang, J. Huang, S. Jiang, L. Du, *Viruses* **11**, 60 (2019).
15. L. Du *et al*., *Expert Opin. Ther. Targets* **21**, 131–143 (2017).
16. L. Du *et al*., *Nat. Rev. Microbiol.* **7**, 226–236 (2009).
17. Q. Wang *et al*., *Nat. Microbiol.* **4**, 1231–1241 (2019).
18. Y. Wu *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7564–E7573 (2017).

ACKNOWLEDGMENTS

We thank the staff of the BL17U1 beamline at the Shanghai Synchrotron Radiation Facility (SSRF) for data collection. We thank H. Sun for discussion of the histopathology experiment.
Funding: This work was supported by the Zhejiang University special scientific research fund for COVID-19 prevention and control (2020XGZX019), the National Science and Technology Major Project (2018ZX10733403), the National Key R&D Program of China (2018YFC1200603), the National Key Plan for Scientific Research and Development of China (2016YFD0500304), National Natural Science Foundation of China (31872745, 81902058, and 32041010), and the National Science and Technology Major Projects of Infectious Disease Funds (2017ZX10304402). **Author contributions:** Y.W., F.G., G.F.G., and L.L. initiated and coordinated the project. Y.W., F.G., and G.F.G. designed the experiments. Y.L. and L.L. provided the convalescent PBMCs. Y.W. and C.S. performed the cell sorting. W.P., C. Zhao, Z.L., and C.S. sequenced and constructed the antibodies. Y.W. conducted the SPR and Octet analysis with the help of Z.F., F.W., Z.L., and Q.W. C.S. and Y.Y. evaluated the neutralizing potency. S.L. and Y.B. conducted the animal experiments with the help from G.W., W.T., X.L., C.F., C. Zhang and H.X. qRT-PCR was conducted by Y.G. F.W., W.P., D.L., C. Zhao, and Z.L. expressed and purified proteins. J.Q. and F.G. collected the diffraction data and determined the complex structure. Y.W., F.G., and G.F.G. analyzed the data and wrote the manuscript. S.T. revised the manuscript. **Competing interests:** Y.W., F.W., C.S., D.L., S.T., Y.L., G.F.G., and L.L. are listed as inventors on pending patent applications for mAb B38 and H4. The other authors declare that they have no competing interests. **Data and materials availability:** The atomic coordinate for the complex of the COVID-19 virus RBD and B38 has been deposited in the Protein Data Bank (www.rcsb.org). The PDB ID is 7BZ5. Antibody sequences beyond those in the supplementary materials are available under an MTA. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/368/6496/1274/suppl/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 to S4
References (*19*–*22*)
MDAR Reproducibility Checklist

13 April 2020; accepted 5 May 2020
Published online 13 May 2020
10.1126/science.abc2241

Multiple Positions Open in The Department of History of Science, Technology and Medicine, Peking University

The Department of History of Science, Technology and Medicine (DHSTM) at Peking University is currently inviting applications for multiple positions at the ranks of full professor with tenure, associate professor with tenure, assistant professor on tenure track, and post-doctoral fellow respectively. Given the department's current engagement in the historical study of science in modern and contemporary China as well as related explorations based on big data, candidates with relevant experience are especially encouraged to apply.

Specific Qualification:

1.Full Professor with tenure/ Associate Professor with tenure/ Assistant Professor on tenure track

Applicants should be highly influential scholars with significant publications in their own field and working experience in a top academic institution, either in China or overseas. Candidates should have rich teaching experience, extensive academic connections, as well as excellent peer review. Preferred candidates are those specializing in the twentieth-century history of science, technology, and Medicine (including both Chinese and Western) or in the philosophy of science.

2.Postdoctoral Fellows

Applicants should have demonstrated outstanding academic potential as well as a doctorate in any of the following specialties: the history of science and technology, the philosophy of science, the STS (science, technology and society), science communication and Big Data research in the history of science, history, or sociology.

Contacts:
Mr. Zheng
Email: hstm@pku.edu.cn
Address: Room 211, Jingyuan Courtyard 1, Peking University, Beijing 100871, China

DEDICATED TEAM FOR JOB CONSULTING

学术桥 AcaBridge.edu.cn

Professional Career Service for Scholars

4 Exclusive Accesses

Unique One-to-one Consultation
Customized Job Map
Full Latest Information
Talent Policy Q&A

Professional Career Guide and Expert Advice
Acabridge is Your Top Choice !

www.acabridge.edu.cn

Overseas Chinese Scholars' Visit to Top Chinese Universities

Check the Details from www.edu.cn/zgx

- 10,000+ academic job vacancies in China
- Free one-to-one consultation service

Send your CV to *consultant@acabridge.edu.cn*

海外学者中国行

中山大學　大连理工大学　杭州師範大學
上海海事大學　上海外国语大学

Scan for application

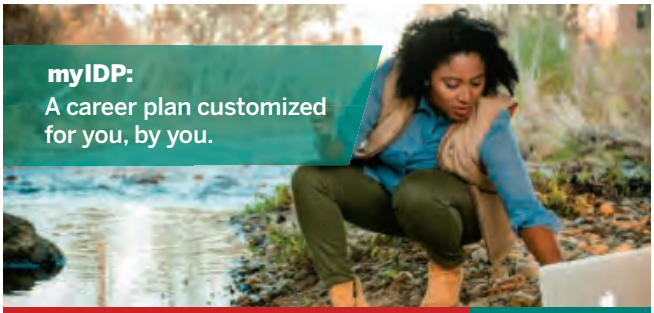
ScienceCareers online @sciencecareers.org

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

Confused about your next career move?

Download Free Career Advice Booklets!

ScienceCareers.org/booklets



For your career in science, there's only one Science

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.

myIDP

Visit the website and start planning today!
myIDP.sciencecareers.org

Science Careers In partnership with:

FASEB · University of Massachusetts Medical School · UCSF University of California San Francisco · UIC · Burroughs Wellcome Fund

EGL Charitable Foundation

The EGL Charitable Foundation invites you to apply to the

Gruss Lipper Post-Doctoral Fellowship Program

Eligibility

- Israeli citizenship
- Candidates must have completed PhD and/or MD/PhD degrees in the Biomedical Sciences at an accredited Israeli University/Medical School or be in their final year of study
- Candidates must have been awarded a postdoctoral position in the U.S. host research institution.

Details regarding the fellowship are available at www.eglcf.org

Application Deadline is October 4, 2020

Who's the Top Employer for 2019?

Science Careers' annual survey reveals the top companies in biotech & pharma voted on by *Science* readers.

Read the article and employer profiles and listen to podcasts at sciencecareers.org/topemployers

By **Moamen M. Elmassry**

Helping others—and myself

During my first year of grad school, a faculty member asked whether I'd be interested in analyzing data for one of his projects. I enjoyed new computational challenges, so I agreed to do the analyses on top of my normal Ph.D. work. Then, 1 year later, the same faculty member met with me and asked a series of questions posed by reviewers of the paper he had written. I was flustered—I had no idea that a manuscript had even been submitted to a journal—but I answered his questions. Later, though, I got up the courage to drop by his office and find out whether I was listed as a co-author, or even acknowledged, on the manuscript. He said no, acting as though the question itself was inappropriate. The experience led me to rethink my approach to collaborations.

I started to develop my computational skills when I was working as a pharmacist in Egypt. I had grown tired of my day job, and I was looking for a new challenge. In my spare time, I took online courses that taught me how to write computer code, and I realized that I enjoy programming. Eventually, I decided to quit my pharmacist job and move to the United States, where I started a Ph.D. program in biology.

In grad school, I continued to develop my computational skills. Colleagues noticed my expertise and started to ask for help. My dad was always lending a hand to those in need when I was growing up, and he taught me to do the same. So during my first years of grad school, I always said yes when I was asked to assist with data analyses.

"If you invest time in a project, then you deserve to have your efforts acknowledged."

Once, a fellow grad student came to me with an urgent request. She was leaving for a conference the following week and hadn't analyzed the data for her poster. "Could you help me?" she asked in a desperate tone. Feeling bad for her, I agreed to do some analyses—which took 2 days of my time.

My friends often criticized me for saying yes to these requests, pointing out I needed to focus on my own work. They also worried that I was giving people the benefit of my expertise without getting anything substantive, such as authorship or acknowledgment, in return. I understood my friends' concerns. The grad student who needed help with her poster said she'd return the favor someday and buy me a meal, but she never did. And the faculty member I did analyses for never did put my name on his manuscript.

In time I realized that although it's OK to lend a helping hand, it's also important to speak up for myself. In the years since, I've made a point of having a conversation about authorship as soon as possible in a collaboration if I'm being asked for a significant contribution to a project. These can be awkward conversations, especially when the collaborator is more senior than I am. But I've learned that they are necessary, because if you invest time in a project, then you deserve to have your efforts acknowledged.

These discussions pay off in other ways, too. Ensuring that everyone involved is on the same page alleviates stress during the collaborative process. That frees me to focus on the science, which is the truly fun part. The conversations have also resulted in accomplishments that I could add to my CV. I defended my Ph.D. 3 months ago, and I left grad school as a co-author on 10 manuscripts—with a few more on the horizon.

The collaborations I participated in during my time as a Ph.D. student enriched my grad school experience, giving me an opportunity to learn about and contribute to projects apart from my own. I particularly enjoyed helping fellow grad students who were interested in honing their quantitative skills. It was rewarding to see their skills blossom, often to the point where they didn't need my help anymore.

But along the way, I learned that in order to maintain my sanity during the collaborative process, I need to be transparent about my expectations. Collaborators can't read my mind, and they may not automatically appreciate how much time it takes to analyze data. So it's my responsibility to tell them. ■

Moamen M. Elmassry is a postdoctoral research associate at Princeton University in New Jersey. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER

CALL FOR PAPERS

SPJ
SCIENCE PARTNER JOURNALS

SIBET CAS

spj.sciencemag.org/bmef

BME Frontiers

OPEN ACCESS

Biomedical Engineering (BME) Frontiers is a **Science Partner Journal** distributed by the **American Association for the Advancement of Science** (AAAS) in collaboration with the **Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences** (SIBET CAS). *BME Frontiers* aims to serve as an effective platform for the multidisciplinary community of biomedical engineering. The journal will publish breakthrough research in the fields of pathogenic mechanisms as well as disease prevention, diagnosis, treatment, and assessment.

The Science Partner Journals (SPJ) program was established by the American Association for the Advancement of Science (AAAS), the nonprofit publisher of the *Science* family of journals. The SPJ program features high-quality, online-only, open access publications produced in collaboration with international research institutions, foundations, funders and societies. Through these collaborations, AAAS expands its efforts to communicate science broadly and for the benefit of all people by providing top-tier international research organizations with the technology, visibility and publishing expertise that AAAS is uniquely positioned to **offer as the world's largest general science membership society**.

Submit your research to *Biomedical Engineering Frontiers* today!
Learn more at: **spj.sciencemag.org/bmef**

ARTICLE PROCESSING CHARGES WAIVED UNTIL 2021

ScienceSignaling.org

PUT YOUR RESEARCH OUT IN FRONT

Submit your research:
cts.ScienceMag.org

Science Signaling
AAAS

Twitter: @SciSignal
Facebook: @ScienceSignaling